

General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.
- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.
- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.
- This document is paginated as submitted by the original source.
- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.



PROCESS AIR QUALITY DATA Final Technical
Report, Jan. 1977 - Mar. 1978 (NASA) 112 p
HC A06/MF A01 CSCL 13B

N79-19476

G3/45 Unclas
16328

FINAL TECHNICAL REPORT

on

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
GRANT NUMBER NSG 1375

for



PROCESSING AIR QUALITY DATA

NASA

FINAL TECHNICAL REPORT

on

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
GRANT NUMBER NSG 1375

for

PROCESSING AIR QUALITY DATA

by

Dr. Charles M. Butler, Professor
and
Dr. Thomas E. Pritchard, Associate Professor
Department of Business Statistics and Data Processing
Mississippi State University
Mississippi State, Mississippi 39762

May 22, 1978

A B S T R A C T

Final Technical Report on NASA Grant NSG 1375 (January 1977 - March 1978),
Processing Air Quality Data: The Virginia State Air Pollution Control Board (VSAPCB) has conducted air quality sampling since January 1968. Data for air quality parameters, recorded on written forms, punched cards or magnetic tape, are available from the VSAPCB Region VI - Hampton Roads for 1972 through 1975. The principal investigator for this research project participated at the Langley Research Center (LRC) as a National Aeronautics and Space Administration - American Society for Engineering Education (NASA-ASEE) 1976 Summer Research Fellow in the initial phase of computer software development by the LRC Analysis and Computation Division (ACD) for the processing of air quality data. During this initial phase, computer software was developed to (1) calculate several daily statistical measures of location, (2) plot time histories of data or the calculated daily statistics, (3) calculate simple correlation coefficients, and (4) plot scatter diagrams. This research project continued at Mississippi State University (MSU) the development of computer software for processing air quality data to include time series analysis and goodness of fit tests. The usual multiplicative time series model with trend, seasonal, cyclical and irregular components, $Y = T \times S \times C \times I$, was extended to include daily and weekly periodic components, $Y = T \times D \times W \times S \times C \times I$. Computer software was developed to (1) calculate a larger number of daily statistical measures of location, (2) calculate a large number of daily, monthly and yearly measures of location, dispersion, skewness and kurtosis, (3) decompose the extended time series model and (4) perform some goodness of fit tests. The computer program is described, documented and illustrated by examples. Recommendations are made for continuation of the development of research on processing air quality data.

PROCESSING AIR QUALITY DATA

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT.	ii
TABLE OF CONTENTS	iii
I. BACKGROUND.	1
II. RESEARCH PROPOSED.	2
A. INTRODUCTION.	2
B. THE TIME SERIES ANALYSIS MODEL.	3
C. GOODNESS OF FIT TESTS	7
III. RESEARCH ACCOMPLISHED	9
A. THE EXTENDED TIME SERIES ANALYSIS MODEL.	9
B. THE COMPUTER PROGRAM	22
IV. RECOMMENDATIONS.	23
V. REFERENCES.	24
<u>ATTACHMENT: COMPUTER PROGRAM USER'S MANUAL</u>	
INTRODUCTION	A 1
USER INPUT CARDS	A 4
SUBROUTINES.	A12
DETAILED DESCRIPTION OF PROGRAMS	A15
APPENDIX	A22

I. BACKGROUND

Air quality data has been collected by the Virginia State Air Pollution Control Board (VSAPCB) since January 1968. Among the seven VSAPCB regions, Region VI - Hampton Roads encompasses the Norfolk-Peninsula area from the North Carolina border to the Williamsburg-Yorktown vicinity. The National Aeronautics and Space Administration (NASA) has entered into a cooperative agreement with the VSAPCB, by and through the Region VI Director, whereby the Langley Research Center (LRC) assists the Region VI office by making available the technological capabilities of the LRC that may be of value to the Region VI office in accomplishing the VSAPCB's monitoring and emissions control programs.

In addition to the considerable expertise of the LRC for assisting the Region VI office of the VSAPCB in the realm of sampling instrumentation, the vast computer complex, computer programming, and mathematical analysis capabilities of the LRC Analysis and Computation Division (ACD) can be of invaluable assistance in the data processing of very large amounts of air quality data collected and stored over the past several years.

The principal investigator (Dr. Charles M. Butler) for this research project had the privilege of participating as a NASA-ASEE (American Society for Engineering Education) Research Fellow in the initial phase of the LRC effort to assist in the analysis and interpretation of Region VI's accumulated air quality data. During this initial phase (Summer 1976), computer software was developed to access the stored data, to calculate some summary statistics and correlation coefficients, and to plot some time histories and scatter diagrams.

The proposed purpose of this research project was to continue the development of computer software for analysis and interpretation of air quality data.

II. RESEARCH PROPOSED

A. INTRODUCTION

Any complete description of data processing must include the five steps:

1. generation,
2. recording,
3. validation,
4. storage and retrieval, and
5. analysis and interpretation¹.

A great deal of effort has been and continues to be expended on the first, second and fourth of these steps. Step three probably needs careful study but was not the subject of this research project. Step five, the analysis and interpretation of air quality data, was the topic to be investigated herewith.

During the summer of 1976, the principal investigator of this proposed research project participated in the initial phase of the LRC-VSAPCB effort to develop software for meaningful analysis and interpretation of air quality data collected in Region VI over the past several years. This data consists of continuous hourly averages of sulphur dioxide (SO₂), carbon monoxide (CO), ozone (O₃), and wind velocity; continuous two-hourly averages of soiling index (coefficient of haze - COH); and daily averages

¹Superscripts refer to numbered references in Section V of this report.

of total suspended particulate (TSP), collected either every third day or every sixth day. Such data may be recorded on written forms, punched cards, or magnetic tape; and it may be in standard or non-standard formats. Since all such data may be transformed to the standard SAROAD formats¹ and may be recorded on magnetic tape, the software was to be developed for accessing data on magnetic tape in the standard SAROAD formats.

As of August 13, 1976, under the supervision of Dr. Stephen K. Park of LRC's ACD, a software package had been developed at LRC¹¹ to plot (1) continuous hourly or two-hourly time histories for each month; (2) daily arithmetic mean, minimum, maximum, midrange and median for each month; and (3) correlation scatter diagrams for pairs of parameters at each site and at pairs of sites for each parameter for each month.

This research project was proposed to continue development of this software package to (1) decompose the data for each parameter into long term trend, seasonal variation, cyclical fluctuation and irregular components and (2) perform goodness of fit tests for various standard distribution models to the air quality data.

This research project was proposed to consist of two parts: time series analysis and goodness of fit tests.

B. THE TIME SERIES ANALYSIS MODEL^{5,6,7,8}

A time series of the type considered here is a sequence of values of some variable (or composite of variables) arranged in the chronological order of their occurrence. Each value (datum) is assumed to consist of four components:

Trend, denoted by T, is the long-term (many years) smooth variation over time. Most often T is considered linear, but is of other forms if

the situation or data so indicates; polynomials, exponentials, and power functions are not uncommon. The method of least squares is usually the method used to fit the trend curve to the data, although experienced practitioners may frequently use visual or other short-cut methods for data with which they are thoroughly familiar.

Seasonal variations, denoted by S, is a conglomerate of all periodic variations in the data whose periods are one year or less. Usually S is considered in the form of an index number which expresses seasonal variations as a percentage of the corresponding trend (or combination of trend and other components) values. There are two basic methods of extracting the seasonal variations component from time series data of the type considered here: the ratio-to-trend and the ratio-to-moving-average methods. Where there is not yet enough experience with a situation or type of data to know which method produces the "best" results, both methods should be applied and their results compared to see which produces the better representation of the data or situation being studied.

Cyclical fluctuations, denoted by C, is a conglomerate of all periodic fluctuations in the data whose periods are greater than one year -- in practice, usually up to about three to seven years, depending on the nature of the data. C is usually expressed as an index number giving the cyclical component as a percentage of trend or combination of trend and seasonal variations. The ratio-to-moving-average method is used to extract the cyclical component in situations where it is believed to be of sufficient importance, but, quite often, it is not considered as a separate entity, being left in combination with S when using the ratio-to-trend method.

Irregular or random movements, denoted by I, consists of movements in data values that are not attributable to trend, seasonal variations

or cyclical fluctuations. There is usually no effort made to analyze the irregular component, but only to recognize it where it occurs in the data and remove or minimize its effect on the data values. Often the I component is attributable to tornado, hurricane, lightning, flood, fire or other rare and unpredictable phenomena.

The time series model consists of the combination of the four components in some multiplicative or additive fashion, although various combinations and transformations may be appropriate in some unusual circumstances. The model most often used is the straightforward multiplicative model $Y = T \times S \times C \times I$, where each datum is considered to be the simple product of the four components.

If the ratio-to-trend method is used, a trend line is fitted to the data, usually by the method of least squares. The trend value corresponding to each datum y-value may then be calculated by substituting that datum x-value into the expression for the trend line, $\hat{T} = a + bX$ (assuming a linear trend -- if other than a linear trend line is used, the same procedure is followed with that fitted trend curve). Each trend value is then divided into its corresponding datum y-value (In terms of the model this is $\frac{Y}{\hat{T}} = S \times C \times I$.) yielding an index number (It is customary to multiply by 100 to discuss the index number in terms of percentages.) which expresses the $S \times C \times I$ component of each datum y-value as a proportion of the corresponding trend value. In the ratio-to-trend method this quotient is referred to as the seasonal index; this is rather "loose" terminology, since it actually is an index of combined seasonal, cyclical and irregular movement. If there is sufficient data (more than one year) for repeated calculations of the same index (e.g., for March 1973, March 1974, and March 1975), the repeated indexes are usually averaged to

produce one index for that point (month, day, hour, week, etc.) in the time series. The ratio-to-trend method then uses the trend line (curve) to forecast or predict other values of the variable being studied, usually in the near future (Of course, the usual precautions about extrapolation are in order.); these forecast or predicted trend values are then corrected for seasonal variation (actually combined seasonal, cyclical, and irregular variations when the ratio-to-trend method is used) to produce the value expected if and when the actual data is collected. (The usefulness of such a forecasting ability for air pollution data would be to cut back on emissions to avoid the forecast value when it is beyond allowable limits.)

If the ratio-to-moving-average method is used, the trend line may be fitted (usually by the least squares method) either before or after the seasonal component is extracted -- or even after extracting both seasonal and cyclical components; it is not clear what effect the order of the operations has on the final decomposition of a time series and a selection is usually made in a previously unexplored situation after comparing the results of the various orders in which the operations may be performed. The seasonal component is extracted by dividing each datum y-value by the annual average centered on that datum. (In terms of the model this is $\frac{Y}{T \times C \times I} = S$.) If it is desired to extract the cyclical component, each datum y-value is divided by a three to seven (or more, depending on the nature of the data) years average centered on that datum (In terms of the model this is $\frac{Y}{T \times I} = S \times C$.) followed by division by the S value found in the preceding annual ratio-to-moving-average calculation ($\frac{S \times C}{S} = C$). Then a trend value is forecast or predicted, corrected for seasonal or seasonal-and-cyclical variation (The cyclical correction is

ignored if the preceding analysis and calculations have indicated it is of little importance.), and used as appropriate in the analysis. (In analysis of air pollution data, this may prove to be a valid method of predicting parameter values that are expected to exceed allowable limits, allowing a reduction in emissions to prevent the undesirable excess.)

C. GOODNESS OF FIT TESTS^{6,7,15}

The two commonly used tests for deciding how well a theoretical distribution fits a set of data reduced to a frequency distribution are the χ^2 (chi-square) and Kolmogorov-Smirnov goodness of fit tests. Both are appropriate for the problem of fitting a theoretical distribution to air pollution data of the type considered here, but only the χ^2 will be described. (The Kolmogorov-Smirnov test could be applied later to see how well it agrees with the χ^2 test.)

The data is first reduced to a frequency distribution using relative frequencies and a number of classes appropriate for the data being studied. (There is no concrete rule but a good rule of thumb is to use between five and fifteen classes -- exceptions to the rule are not uncommon and only a trial-and-error approach can decide the number of classes to be used in a previously unexplored situation.)

A theoretical distribution is selected as a possible source of the data being studied, e.g., for total suspended particulate, the literature indicates a lognormal distribution⁴ has been used by most recent practitioners. (I believe a study should be made, mostly empirical, testing the goodness of fit of a variety of distributions--including some modification of the Poisson--to decide what theoretical distribution to use for each of the air pollution parameters studied.) The probability for each

of the frequency distribution class intervals is obtained for the selected distribution (e.g., lognormal for total suspended particulate) using parameter estimates calculated from the sample data (frequency distribution).

For example, if $Y = e^X$ is lognormal, X is normal with mean μ and variance σ^2 , and μ and σ^2 are considered to be the parameters of Y also. Then the mean of Y is $\mu_Y = e^{\mu + \frac{1}{2}\sigma^2}$ and the variance of Y is $\sigma_Y^2 = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$.

The parameters would be estimated by $\mu \doteq \bar{X} = \frac{1}{n} \sum_{i=1}^n \ln Y_i / n \doteq \bar{X}' = \frac{1}{c} \sum_{i=1}^c f_i \cdot \ln Y_i' / n$ and by $\sigma^2 \doteq s^2 = \frac{1}{n} \sum_{i=1}^n (\ln Y_i - \bar{X})^2 / (n - 1) \doteq \frac{1}{c} \sum_{i=1}^c (\ln Y_i' - \bar{X}')^2 / (n - 1)$, where $Y_i = e^{X_i}$ for $i = 1, 2, \dots, n$ are the n data values; or $Y_i' = e^{X_i'}$ for $i = 1, 2, \dots, c$ are the c class marks and f_i for $i = 1, 2, \dots, c$ are the c class frequencies; and $\ln$ is an abbreviation for natural logarithm. (Since two values have been calculated from the data to obtain the theoretical frequencies, the number of degrees of freedom for the χ^2 test will be $n - 2$.)

A standard normal distribution table is then used to obtain the theoretical frequencies (probabilities) for the classes in the frequency distribution, e.g., if f_{oj} , $j = 1, 2, \dots, c$ is the observed frequency of the j th class which has lower real limit X_{j-1} and upper real limit X_j , the corresponding expected theoretical frequency for that class will be given by $f_{ej} =$

$$P(Y_{j-1} < Y \leq Y_j) = P(Y_{j-1} < e^X \leq Y_j) = P(\ln Y_{j-1} < X \leq \ln Y_j) = P\{(\ln Y_{j-1} - \mu)/\sigma < (X - \mu)/\sigma \leq (\ln Y_j - \mu)/\sigma\} \doteq \Phi\{(\ln Y_j - \bar{X})/s\} - \Phi\{(\ln Y_{j-1} - \bar{X})/s\},$$

using $\mu \doteq \bar{X}$ and $\sigma^2 \doteq s^2$, where Φ is the cumulative normal distribution function. (Hence, any standard normal distribution table can be used to calculate lognormal probabilities. Of course, we are assuming very large sample sizes since we are treating $(X - \bar{X})/s$ as though it were normally distributed.)

A value of χ_o^2 (observed) is then calculated from

$$\chi_o^2 = \sum_{j=1}^c (f_{oj} - f_{ej})^2 / f_{ej}.$$

This value is then compared with tabulated values of χ_v^2 with $v = n - 2$ degrees of freedom to determine at what level of significance the hypothesis that Y is a lognormal distribution may be rejected.

III. RESEARCH ACCOMPLISHED

A. THE EXTENDED TIME SERIES ANALYSIS MODEL

After seeing a few plots of air pollution data as time series of hourly or two-hourly averages, it is possible to surmise periodic components which would not be discernible when using the usual decomposition of the multiplicative time series model $Y = T \times S \times C \times I$. However, it is not difficult to extend this model to incorporate additional multiplicative components with periods of any desirable duration.

The extended time series model consists of the usual four components, T, S, C, and I, as described in I.B. above, together with the following two components:

Daily variations, denoted by D, is a conglomerate of all periodic variations in the data whose periods are 24 hours or less. Like S and C, D will be expressed in the form of an index number.

Weekly variations, denoted by W, is a conglomerate of all periodic variations in the data whose periods are seven days or less. Like S, C, and D, W will be expressed in the form of an index number.

The extended time series model, as considered here, will be expressed as the multiplicative model given by $Y = T \times D \times W \times S \times C \times I$, and the ratio-to-moving-average method will be used to extract first D, and then W components.

To extract D, centered moving averages using 24 hour periods will be calculated for all possible 24 hour periods. This, in effect, eliminates or "smoothes out" all components with periods of 24 hours or less, leaving each centered 24 hour average consisting of the components $T \times W \times S \times C \times I$. Thus, if we divide each of the original $\hat{Y}$ data values by its corresponding centered moving average, the result will be

$$\frac{\hat{Y}}{T \times W \times S \times C \times I} = \frac{T \times D \times W \times S \times C \times I}{T \times W \times S \times C \times I} = \hat{D},$$

which then is the desired index number for the daily variation. We would then average the $\hat{D}$ values for corresponding time periods each day, e.g., midnight to 0200 for two-hourly averages, and use this average- $\hat{D}$ to correct the forecast trend value for that time period. We do likewise for each time period each day.

To extract W, centered moving averages using seven day periods will be calculated, which eliminates or "smoothes out" all components with periods of seven days or less. Now, when we divide each $\hat{Y}$ data value by its corresponding centered seven day moving average, we have

$$\frac{\hat{Y}}{T \times S \times C \times I} = \frac{T \times D \times W \times S \times C \times I}{T \times S \times C \times I} = \hat{D} \times \hat{W}.$$

This result, in turn is divided by the average of the $\hat{D}$ values during that seven day period to get

$$\frac{\hat{D} \times \hat{W}}{\hat{D}_{\text{avg.}}} = \hat{W},$$

which is the desired index number. Again we average the $\hat{W}$ values for each day of the week, and use that average- $\hat{W}$ to correct the trend forecast values which have already been corrected for the average- $\hat{D}$.

We may continue this process for S and C, in turn.

This data is plotted (and connected by straight line segments for clarity) in Figure 1. The first 32 characters in the records above are explained in the upper left corner of Figure 1, in accordance with the standard SAROAD format^{1,2}. The last 48 characters in the records above give the data which is shown again in Table 1.

A linear trend line was fitted to the data with the following results:

$$\hat{T} = a + bX = .90 + .00X = .90.$$

The 24-hour centered moving averages were then calculated, are shown in Table 2 and are plotted in green on Figure 1. The fitting of a linear trend line by the least squares method to the 24-hour centered moving averages yielded:

$$\hat{T}_1 = a_1 + b_1 X_1 = .90 + .00 X_1 = .90.$$

Table 3 shows $\hat{D}$, the ratio of each two-hour average to its corresponding 24-hour centered moving average. The values of $\hat{D}$ were averaged for each period of each day of the week, as shown in Table 4. Each average- $\hat{D}$ value was then applied to its corresponding trend value to produce the forecast shown in Table 5 and plotted in green on Figure 1. The time period average is plotted in Figure 2 for each day of the week and for the overall average.

Next, the seven-day centered moving averages were calculated, are shown in Table 6 and are plotted in red on Figure 1. A linear trend fitted by the least squares method yielded:

$$\hat{T}_2 = a_2 + b_2 X_2 = 1.00 + .00 X_2 = 1.00.$$

Table 7 contains $\hat{D} \times \hat{W}$, the ratio of each two-hour average to its corresponding seven-day centered moving average. The values of $\hat{D} \times \hat{W}$ were averaged for each period of each day, as shown in Table 8. Each average- $\hat{D} \times \hat{W}$ was then applied to its corresponding trend value to produce the forecast shown in Table 9 and plotted in red on Figure 1. The daily average- $\hat{D} \times \hat{W}$ is plotted in Figure 3.

No attempt is made to draw inferences from this one, short-term example. The only purpose here is to illustrate the procedure which is to be programmed. Only after a large number of such examples have been studied in detail should one attempt to draw conclusions about the data using this procedure.

RECORD:

Type of data: 1 (less than 24-hour averages)

State: 48 (Virginia)

Area: 0710 (Chesapeake-city)

Site: 001 (Armory)

Agency: T (State)

Project: 02 (Source-oriented ambient surveillance)

Time: 2 (two-hour averages)

Year: 73 (1973)

Month: 01 (January)

Day: 01 (first day of month)

Starting Hour: 00 (midnight)

Parameter Code: 11201 (soiling index)

Method: 81 (tape sampler transmittance)

Units: 09 (COHS per 1000 linear feet)

Decimal Point: 2 (places or digits to right of decimal point)

Reading 1: 0000 (no data)

Reading 2: 9999 (no data)

Reading 12: 9999 (no data)

NEXT

RECORD:

FOLDED FRAME

ORIGINAL PAGE IS
OF POOR QUALITY

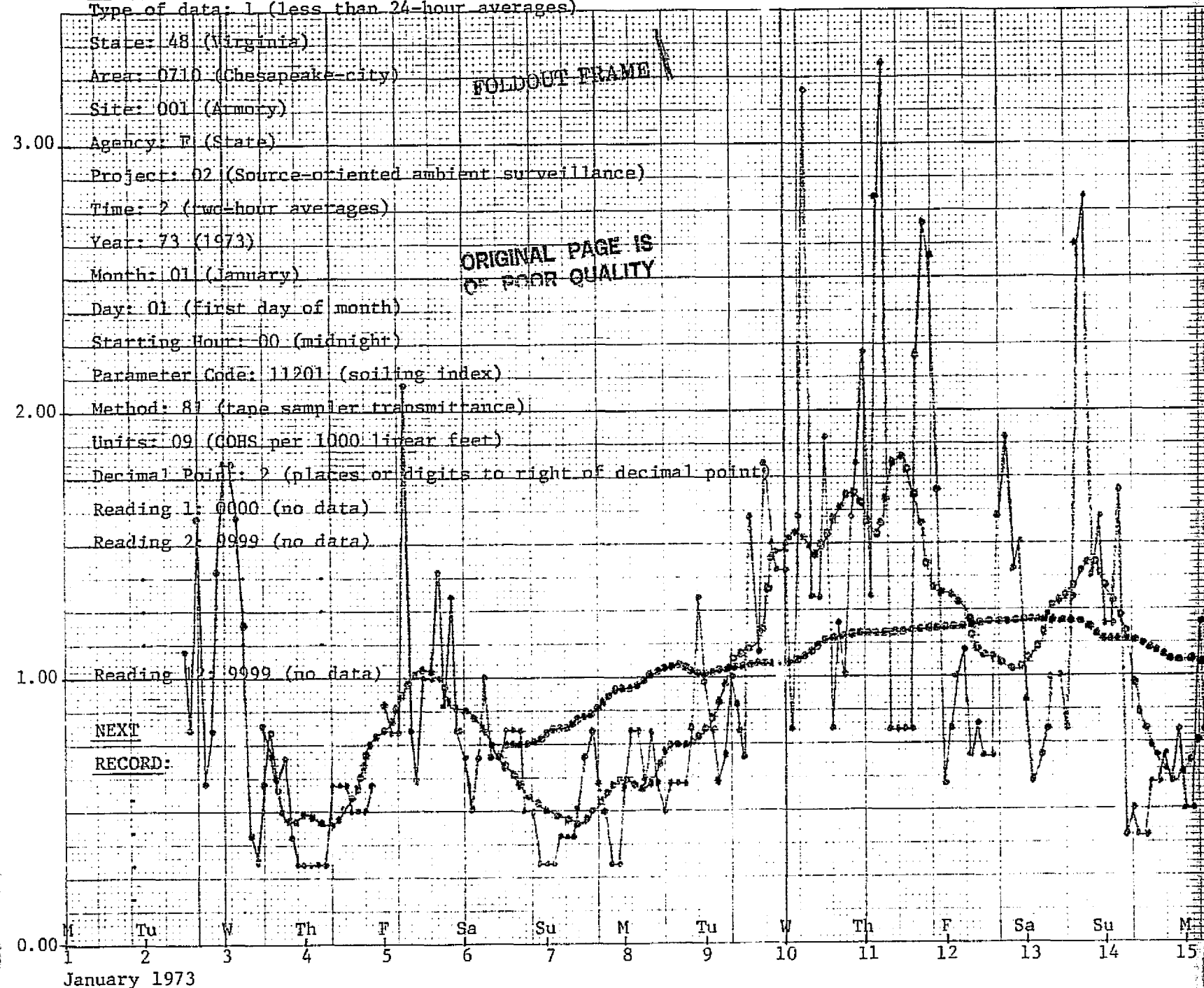
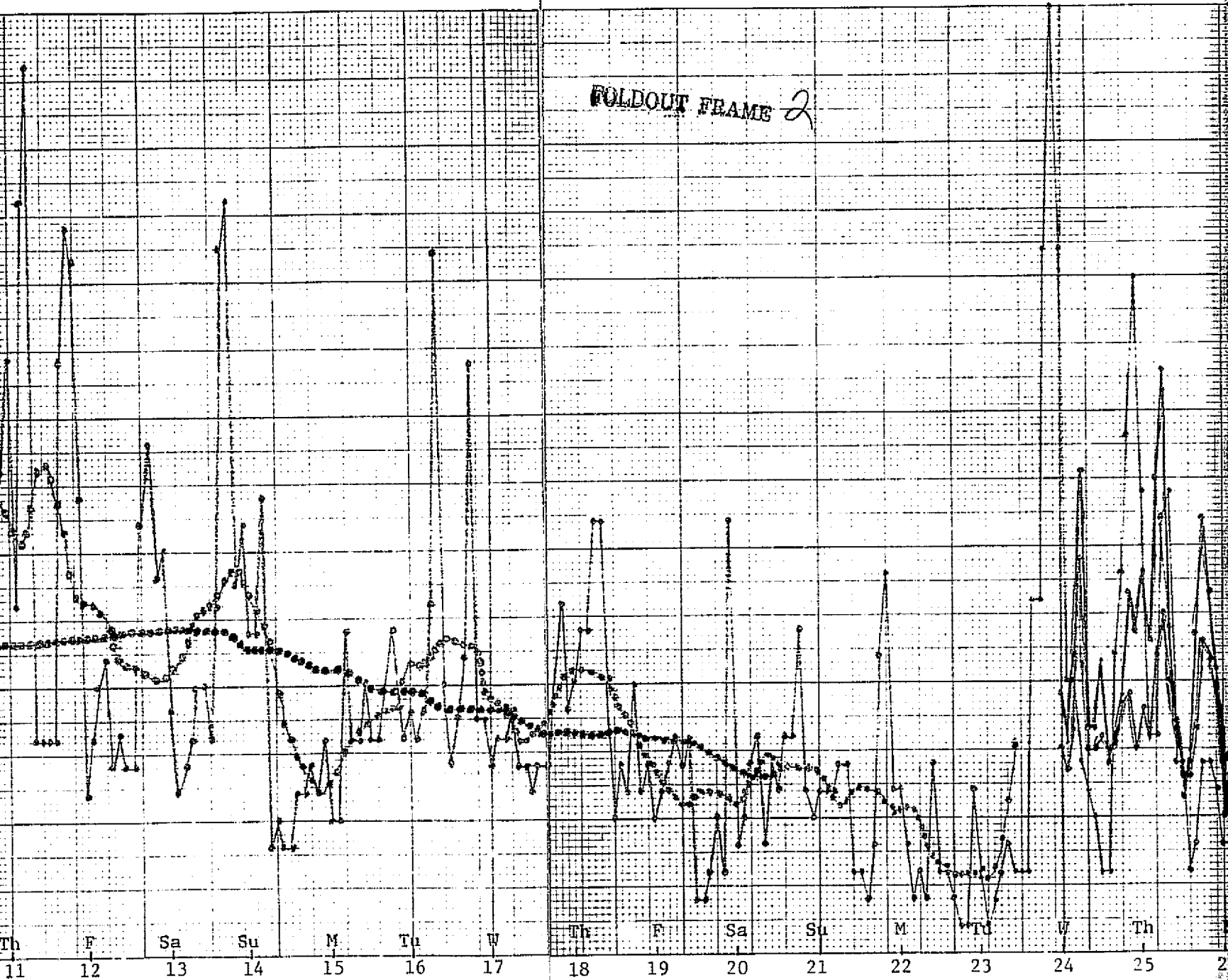


Figure 1

FOLDOUT FRAME 2



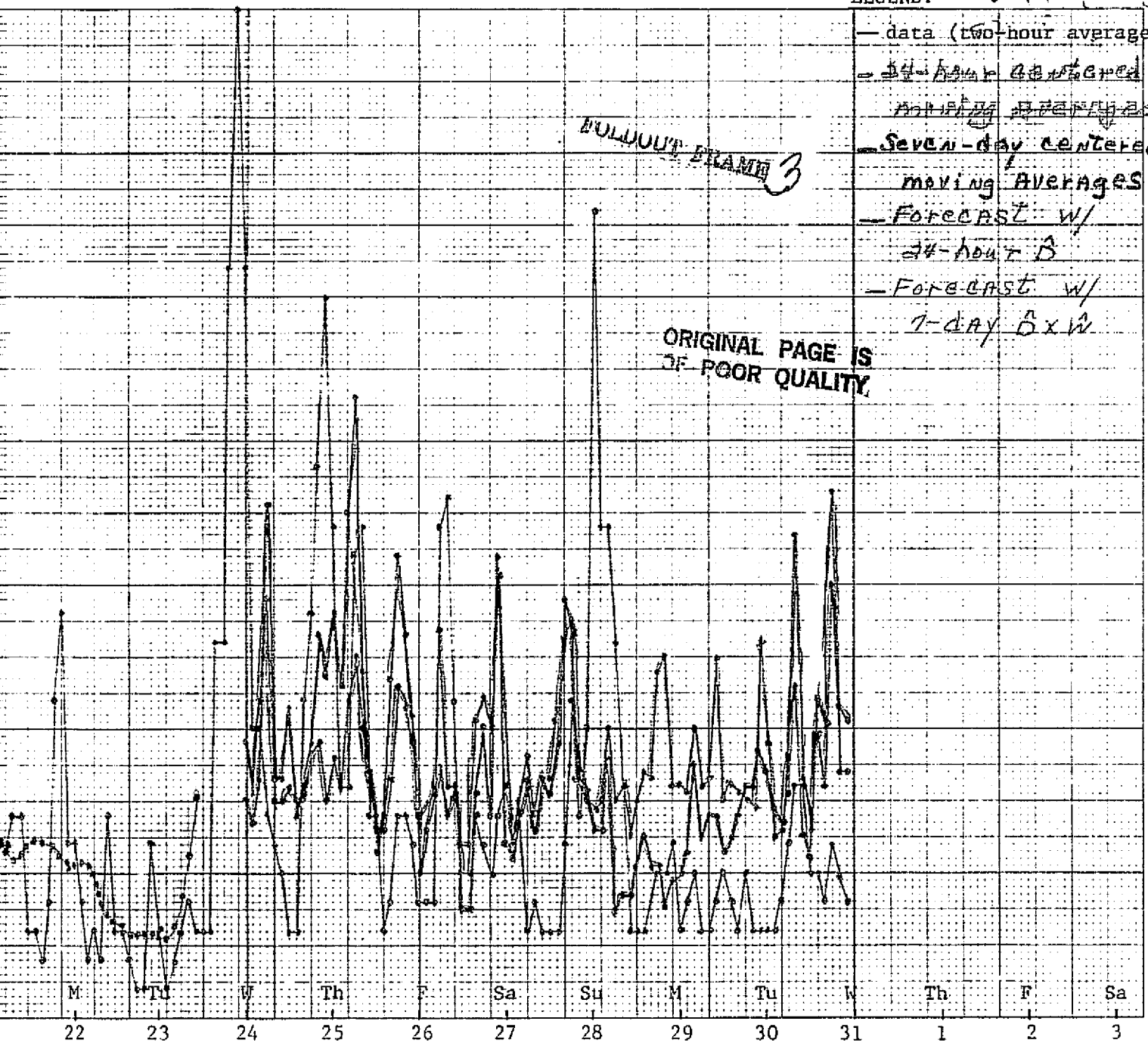
Th 11 F 12 Sa 13 Su 14 M 15 Tu 16 W 17 Th 18 F 19 Sa 20 Su 21 M 22 Tu 23 W 24 Th 25

LEGEND:

- data (two-hour averages)
- 24-hour centered moving averages
- Seven-day centered moving averages
- Forecast w/ 24-hour $\hat{\sigma}$
- Forecast w/ 7-day $\hat{\sigma} \times \hat{w}$

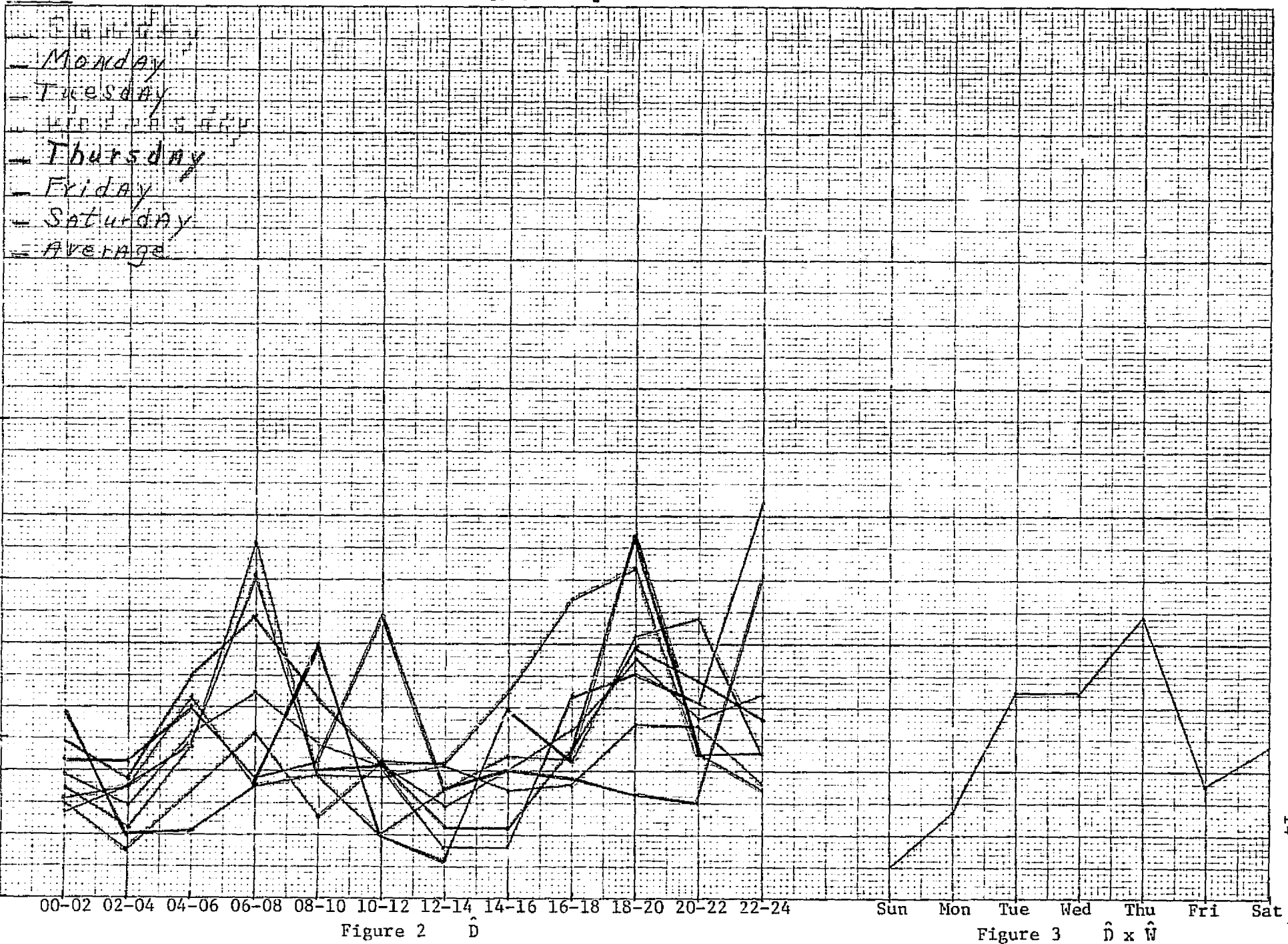
EXHIBIT FRAME 3

ORIGINAL PAGE IS
OF POOR QUALITY



ORIGINAL PAGE IS
OF POOR QUALITY

LEGEND:



Reading and Time Period:												
1973 Day:	1 00-02	2 02-04	3 04-06	4 06-08	5 08-10	6 10-12	7 12-14	8 14-16	9 16-18	10 18-20	11 20-22	12 22-24
Wed Jan 3	01.80	01.80	01.60	01.20	00.40	00.30	00.60	00.80	00.60	00.70	00.40	00.30
Thu Jan 4	00.30	00.30	00.30	00.30	00.60	00.60	00.60	00.50	00.50	00.50	00.60	00.70
Fri Jan 5	00.90	00.80	00.80	02.10	00.80	00.60	01.00	01.00	01.40	00.90	01.30	00.80
Sat Jan 6	00.70	00.50	00.70	01.00	00.70	00.70	00.80	00.80	00.80	00.50	00.50	00.30
Sun Jan 7	00.30	00.30	00.40	00.40	00.40	00.50	00.70	00.80	00.60	00.50	00.30	00.30
Mon Jan 8	00.60	00.80	00.80	00.60	00.80	00.60	00.50	00.60	00.60	00.60	00.80	01.30
Tue Jan 9	01.00	00.80	00.60	00.70	00.10	00.90	00.70	01.60	01.10	01.80	01.50	01.40
Wed Jan 10	01.40	00.90	01.60	03.20	01.30	01.30	01.90	00.80	01.20	01.00	01.60	01.80
Thu Jan 11	02.20	01.30	02.80	03.30	00.80	00.80	00.80	00.80	02.20	02.70	02.60	01.70
Fri Jan 12	00.60	00.80	01.00	01.10	00.70	00.80	00.70	00.70	01.60	01.90	01.40	01.50
Sat Jan 13	00.90	00.60	00.70	00.80	01.00	01.00	00.80	01.30	02.60	02.80	01.40	01.60
Sun Jan 14	01.20	01.20	01.70	00.40	00.50	00.40	00.40	00.60	00.60	00.70	00.60	00.80
Mon Jan 15	00.50	00.50	01.20	00.80	00.80	01.00	00.80	00.80	01.00	01.20	01.00	00.80
Tue Jan 16	00.90	00.80	00.90	01.30	02.60	01.00	00.70	00.90	01.10	02.20	00.90	00.90
Wed Jan 17	00.70	00.80	00.80	00.90	00.70	00.70	00.60	00.70	00.70	01.00	01.30	00.90
Thu Jan 18	01.00	01.20	01.20	01.60	01.60	01.00	00.50	00.70	00.60	01.00	00.60	00.70
Fri Jan 19	00.50	00.60	00.70	00.80	00.70	00.80	00.20	00.20	00.30	00.50	00.30	01.60
Sat Jan 20	00.40	00.50	00.70	00.80	00.40	00.70	00.60	00.80	00.80	01.20	00.60	00.50
Sun Jan 21	00.60	00.60	00.60	00.70	00.70	00.30	00.30	00.20	00.40	01.10	01.40	00.60
Mon Jan 22	00.60	00.40	00.20	00.30	00.20	00.70	00.30	00.30	00.20	00.10	00.10	00.60
Tue Jan 23	00.30	00.10	00.20	00.30	00.40	00.30	00.30	00.30	01.30	01.30	02.60	03.50

Table 1: Data (two-hour averages)

Reading and Time Period:												
1973 Day:	1 00-02	2 02-04	3 04-06	4 06-08	5 08-10	6 10-12	7 12-14	8 14-16	9 16-18	10 18-20	11 20-22	12 22-24
Wed Jan 3							.8125	.6875	.5708	.4792	.4500	.4708
Thu Jan 4	.4833	.4708	.4542	.4417	.4417	.4667	.5083	.5542	.5958	.6917	.7750	.7833
Fri Jan 5	.8000	.8375	.8958	.9500	.9958	1.0292	1.0250	1.0042	.9875	.9375	.8875	.8875
Sat Jan 6	.8834	.8667	.8333	.7916	.7417	.6875	.6500	.6250	.6042	.5667	.5292	.5084
Sun Jan 7	.4958	.4917	.4834	.4750	.4667	.4583	.4708	.5042	.5417	.5667	.5917	.6125
Mon Jan 8	.6084	.5917	.5833	.5875	.6125	.6750	.7334	.7500	.7417	.7375	.7500	.7708
Tue Jan 9	.7916	.8417	.9042	.9750	1.0542	1.0875	1.1084	1.1292	1.1750	1.3208	1.4375	1.4667
Wed Jan 10	1.5333	1.5500	1.5208	1.4917	1.4625	1.4834	1.5333	1.5833	1.6500	1.7042	1.6875	1.6458
Thu Jan 11	1.5792	1.5333	1.5750	1.6875	1.8000	1.8375	1.7667	1.6792	1.5833	1.4166	1.3208	1.3167
Fri Jan 12	1.3125	1.3042	1.2750	1.2167	1.1333	1.0750	1.0792	1.0834	1.0625	1.0375	1.0375	1.0583
Sat Jan 13	1.0708	1.1000	1.1667	1.2458	1.2833	1.2875	1.3042	1.3417	1.4084	1.4333	1.3958	1.3500
Sun Jan 14	1.3083	1.2625	1.1500	.9792	.8584	.7917	.7292	.6708	.6208	.6167	.6458	.6833
Mon Jan 15	.7250	.7500	.7750	.8125	.8500	.8667	.8834	.9125	.9125	.9208	1.0167	1.0917
Tue Jan 16	1.0875	1.0875	1.0958	1.1417	1.1792	1.1792	1.1750	1.1667	1.1625	1.1438	1.0458	.9542
Wed Jan 17	.9375	.9250	.9000	.8333	.8000	.8167	.8292	.8584	.8917	.9375	1.0042	1.0542
Thu Jan 18	1.0625	1.0583	1.0542	1.0500	1.0208	.9834	.9542	.9083	.8625	.8084	.7375	.6917
Fri Jan 19	.6708	.6375	.6042	.5708	.5375	.5625	.5958	.5875	.5833	.5833	.5708	.5542
Sat Jan 20	.5667	.6083	.6542	.7042	.7458	.7125	.6750	.6875	.6875	.6792	.6875	.6833
Sun Jan 21	.6542	.6167	.5750	.5542	.5833	.6208	.6250	.6167	.5916	.5583	.5208	.5167
Mon Jan 22	.5333	.5375	.5334	.4833	.3875	.3333	.3208	.2958	.2833	.2833	.2916	.2833
Tue Jan 23	.2667	.2667	.3125	.4083	.5625	.7875						

Table 2: 24-hour Centered Moving Averages

1973 Day:	Reading and Time Period;											
	1 00-02	2 02-04	3 04-06	4 06-08	5 08-10	6 10-12	7 12-14	8 14-16	9 16-18	10 18-20	11 20-22	12 22-24
Wed Jan 3							.7385	1.1636	1.0512	1.4608	.8889	.6372
Thu Jan 4	.6207	.6372	.6605	.6792	1.3584	1.2856	1.1804	.9022	.8392	.7228	.7742	.8936
Fri Jan 5	1.1250	.9552	.9600	2.2105	.8034	.5830	.9756	.9958	1.4177	.9600	1.4648	.9014
Sat Jan 6	.7924	.5769	.8400	1.2633	.9438	1.0182	1.2308	1.2800	1.3241	.8823	.9448	.5901
Sun Jan 7	.6051	.6101	.8275	.8421	.8571	1.0910	1.4868	1.5867	1.1076	.8823	.5070	.4898
Mon Jan 8	.9862	1.3520	1.3715	1.0213	1.3061	.8889	.6818	.8000	.8090	.8140	1.0667	1.6866
Tue Jan 9	1.2633	.9505	.6636	.7179	.9486	.8276	.6315	1.4169	.9362	1.3628	1.0435	.9545
Wed Jan 10	.9131	.5806	1.0521	2.1452	.8889	.8764	1.2392	.5053	.7273	.5868	.9481	1.0937
Thu Jan 11	1.3931	.8478	1.7778	1.9556	.4444	.4354	.4528	.4764	1.3895	1.9060	1.9686	1.2911
Fri Jan 12	.4571	.6134	.7843	.9041	.6177	.7442	.6486	.6461	1.5059	1.8313	1.3494	1.4174
Sat Jan 13	.8405	.5454	.6000	.6422	.7792	.7767	.6134	.9689	1.8461	1.9535	1.0030	1.1852
Sun Jan 14	.9172	.9505	1.4783	.4085	.5825	.5052	.5485	.8944	.9665	1.1351	.9291	1.1708
Mon Jan 15	.6896	.6667	1.5484	.9846	.9412	1.1538	.9056	.8767	1.0959	1.3032	.9836	.7328
Tue Jan 16	.8276	.7356	.8213	1.1386	2.2049	.8480	.5957	.7714	.9462	1.9234	.8606	.9432
Wed Jan 17	.7467	.8649	.8889	1.0800	.8750	.8571	.7236	.8155	.7850	1.0667	1.2946	.8537
Thu Jan 18	.9412	1.1339	1.1383	1.5238	1.5674	1.0169	.5240	.7707	.6956	1.2370	.8136	1.0120
Fri Jan 19	.7454	.9412	1.1586	1.4015	1.3023	1.4222	.3357	.3404	.5143	.8572	.5256	2.8870
Sat Jan 20	.7058	.8220	1.0700	1.1360	.5363	.9825	.8889	1.1636	1.1636	1.7668	.8727	.7317
Sun Jan 21	.9172	.9729	1.0435	1.2631	1.2001	.4832	.4800	.3243	.6761	1.9703	2.6882	1.1612
Mon Jan 22	1.1251	.7442	.3750	.6207	.5161	2.1002	.9352	1.0142	.7060	.3530	.3429	2.1179
Tue Jan 23	1.1248	.3750	.6400	.7348	.7111	.3810						
Average $\bar{D}$	.8868	.7938	.9850	1.1336	.9692	.9138	.7908	.8856	1.0252	1.2488	1.0635	1.1375

Table 3: Ratio of two-hour Averages to 24-hour Centered Moving Averages

Reading and Time Period:													
Day:	1 00-02	2 02-04	3 04-06	4 06-08	5 08-10	6 10-12	7 12-14	8 14-16	9 16-18	10 18-20	11 20-22	12 22-24	Daily Avg.
Sat:	.7796	.6481	.8367	1.0139	.7531	.9258	.9110	1.1375	1.4446	1.5342	.9402	.8357	.9800
Sun:	.8132	.8445	1.1164	.8379	.8799	.6931	.8384	.9351	.9167	1.3292	1.3748	.9406	.9600
Mon:	.9336	.9210	1.0983	.8755	.9211	1.3810	.8409	.8970	.8703	.8234	.7977	1.5124	.9894
Tue:	1.0719	.6870	.7083	.8638	1.2882	.6855	.6136	1.0942	.9412	1.6431	.9520	.9488	.9581
Wed:	.8299	.7228	.9705	1.6126	.8820	.8668	.9004	.8281	.8545	1.0381	1.0439	.8615	.9509
Thu:	.9850	.8744	1.1922	1.3862	1.1234	.9126	.7191	.7164	.9748	1.2886	1.1855	1.0656	1.0353
Fri:	.7758	.8366	.9676	1.5054	.9078	.9165	.6533	.6608	1.1460	1.2162	1.1133	1.7353	1.0362

Table 4: Daily Average-D

Reading and Time Period:												
Day:	1 00-02	2 02-04	3 04-06	4 06-08	5 08-10	6 10-12	7 12-14	8 14-16	9 16-18	10 18-20	11 20-22	12 22-24
Sat:	.70	.58	.75	.91	.68	.83	.82	1.02	1.30	1.38	.85	.75
Sun:	.73	.76	1.00	.75	.79	.62	.75	.84	.82	1.20	1.24	.85
Mon:	.84	.83	.99	.79	.83	1.24	.76	.81	.78	.74	.72	1.36
Tue:	.96	.62	.64	.78	1.16	.62	.55	.98	.85	1.49	.86	.85
Wed:	.75	.65	.87	1.45	.79	.78	.81	.74	.77	.93	.94	.78
Thu:	.89	.79	1.07	1.25	1.01	.82	.65	.64	.88	1.16	1.07	.96
Fri:	.70	.75	.87	1.35	.82	.82	.59	.59	1.03	1.09	1.00	1.56

Table 5: Forecast

Reading and Time Period:												
1973 Day:	1 00-02	2 02-04	3 04-06	4 06-08	5 08-10	6 10-12	7 12-14	8 14-16	9 16-18	10 18-20	11 20-22	12 22-24
Wed Jan 3												
Thu Jan 4												
Fri Jan 5												
Sat Jan 6							.7583	.7506	.7452	.7571	.7744	.7857
Sun Jan 7	.7994	.8071	.8107	.8161	.8250	.8613	.8556	.8786	.8994	.9322	.9512	.9536
Mon Jan 8	.9560	.9589	.9708	.9941	1.0190	1.0369	1.0410	1.0393	1.0405	1.0358	1.0292	1.0298
Tue Jan 9	1.0292	1.0256	1.0250	1.0321	1.0387	1.0434	1.0488	1.0506	1.0512	1.0500	1.0506	1.0542
Wed Jan 10	1.0560	1.0590	1.0726	1.0970	1.1161	1.1292	1.1422	1.1530	1.1660	1.1738	1.1744	1.1744
Thu Jan 11	1.1720	1.1690	1.1678	1.1690	1.1720	1.1768	1.1792	1.1768	1.1777	1.1810	1.1821	1.1845
Fri Jan 12	1.1887	1.1917	1.1952	1.2012	1.2078	1.2041	1.2006	1.2000	1.2018	1.2072	1.2202	1.2305
Sat Jan 13	1.2310	1.2268	1.2226	1.2250	1.2238	1.2172	1.2101	1.2054	1.2000	1.1815	1.1642	1.1571
Sun Jan 14	1.1458	1.1452	1.1345	1.1310	1.1292	1.1220	1.1095	1.1018	1.0917	1.0720	1.0667	1.0726
Mon Jan 15	1.0720	1.0696	1.0595	1.0399	1.0179	1.0000	.9934	.9916	.9887	.9851	.9833	.9833
Tue Jan 16	.9803	.9744	.9637	.9476	.9327	.9268	.9244	.9208	.9202	.9202	.9166	.9113
Wed Jan 17	.9083	.9041	.8905	.8702	.8559	.8446	.8345	.8274	.8172	.8125	.8155	.8161
Thu Jan 18	.8149	.8119	.8083	.8095	.8167	.8208	.8196	.8196	.8131	.8041	.7976	.7922
Fri Jan 19	.7875	.7815	.7738	.7625	.7506	.7440	.7392	.7315	.7232	.7131	.6940	.6768
Sat Jan 20	.6702	.6642	.6619	.6578	.6625	.6881						
Sun Jan 21												
Mon Jan 22												
Tue Jan 23												

Table 6: Seven-day Centered Moving Averages

Reading and Time Period:												
1973 Day:	1 00-02	2 02-04	3 04-06	4 06-08	5 08-10	6 10-12	7 12-14	8 14-16	9 16-18	10 18-20	11 20-22	12 22-24
Wed Jan 3												
Thu Jan 4												
Fri Jan 5												
Sat Jan 6							1.0550	1.0658	1.0735	.6604	.6459	.3818
Sun Jan 7	.3753	.3717	.4934	.4901	.4848	.5805	.8181	.9105	.6671	.5364	.3154	.3146
Mon Jan 8	.6276	.8343	.8240	.6036	.7851	.5786	.4803	.5773	.5766	.5793	.7773	1.2624
Tue Jan 9	.9716	.7800	.5854	.6782	.9627	.8626	.6674	1.5229	1.0464	1.7143	1.4278	1.3280
Wed Jan 10	.3258	.8498	1.4917	2.9170	1.1648	1.1512	1.6624	.6938	1.0292	.8519	1.3624	1.5327
Thu Jan 11	1.8771	1.1121	2.3977	2.8229	.6826	.6798	.6784	.6798	1.8680	2.2862	2.1995	1.4352
Fri Jan 12	.5048	.6713	.8367	.9158	.5796	.6644	.5830	.5833	1.3313	1.5739	1.1474	1.2190
Sat Jan 13	.7311	.4891	.5726	.6531	.8171	.8216	.6611	1.0785	2.1667	2.3699	1.2025	1.3828
Sun Jan 14	1.0473	1.0478	1.4984	.3537	.4428	.3565	.3605	.5446	.5496	.6530	.5625	.7458
Mon Jan 15	.4664	.4675	1.1326	.7693	.7859	1.0000	.8053	.8068	1.0114	1.2182	1.0170	.8136
Tue Jan 16	.9181	.8210	.9339	1.3719	2.7876	1.0790	.7572	.9774	1.1954	2.3908	.9819	.9876
Wed Jan 17	.7707	.8848	.8984	1.0342	.8178	.8288	.7190	.8460	.8566	1.2308	1.5941	1.1028
Thu Jan 18	1.2271	1.4780	1.4846	1.9765	1.9591	1.2183	.6100	.8541	.7379	1.2436	.7522	.8836
Fri Jan 19	.6349	.7678	.9046	1.0492	.9326	1.0753	.2706	.2734	.4148	.7012	.4323	2.3641
Sat Jan 20	.5968	.7528	1.0576	1.2162	.6038	1.0173						
Sun Jan 21												
Mon Jan 22												
Tue Jan 23												

Table 7: Ratio of Two-hour Averages to Seven-day Centered Moving Averages

	Reading and Time Period:												
Day:	1 00-02	2 02-04	3 04-06	4 06-08	5 08-10	6 10-12	7 12-14	8 14-16	9 16-18	10 18-20	11 20-22	12 22-24	Avg. DxW
Sat:	.6640	.6210	.8151	.9346	.7104	.9194	.8580	1.0722	1.6201	1.5152	.9242	.8823	.9614
Sun:	.7113	.7098	.9959	.4219	.4638	.4685	.5893	.7276	.6084	.5947	.4390	.5302	.6050
Mon:	.5470	.6509	.9783	.6864	.7855	.7893	.6428	.6920	.7940	.8988	.8972	1.0380	.7834
Tue:	.9448	.8005	.7596	1.0250	1.8752	.9708	.7123	1.2502	1.1209	2.0526	1.2048	1.1578	1.1562
Wed:	1.0482	.8673	1.1950	1.9756	.9913	.9900	1.1912	.7699	.9492	1.0414	1.4782	1.3178	1.1507
Thu:	1.5521	1.2950	1.9412	2.3997	1.3208	.9490	.6442	.7670	1.3030	1.7649	1.4758	1.1594	1.3810
Fri:	.5698	.7196	.8706	.9825	.7561	.8698	.4268	.4284	.8730	1.1376	.7898	1.7916	.8513

Table 8: Daily Average- $\hat{Dx}\hat{W}$'s

Reading and Time Period:												
Day:	1 00-02	2 02-04	3 04-06	4 06-08	5 08-10	6 10-12	7 12-14	8 14-16	9 16-18	10 18-20	11 20-22	12 22-24
Sat:	.60	.56	.73	.84	.64	.83	.77	.96	1.46	1.36	.83	.79
Sun:	.64	.64	.90	.38	.42	.42	.53	.65	.55	.54	.40	.48
Mon:	.49	.58	.88	.62	.71	.71	.58	.62	.71	.81	.81	.93
Tue:	.85	.72	.68	.92	1.69	.87	.64	1.12	1.01	1.85	1.08	1.04
Wed:	.94	.78	1.08	1.78	.89	.89	1.07	.69	.85	.94	1.33	1.18
Thu:	1.40	1.16	1.75	2.16	1.19	.85	.58	.69	1.17	1.59	1.33	1.04
Fri:	.51	.65	.78	.88	.68	.78	.38	.38	.78	1.02	.71	1.61

Table 9: Forecast

B. THE COMPUTER PROGRAM

A computer program has been developed by Dr. Thomas E. Pritchard, co-principal investigator for this research project, for

- (1) calculation of a large number of measures of location, dispersion, skewness and kurtosis;
- (2) decomposition of an extended time series analysis model; and
- (3) some χ^2 goodness of fit tests.

This computer program is described, documented and illustrated by examples in the attached COMPUTER PROGRAM USER'S MANUAL.

IV. RECOMMENDATIONS

It has become apparent that missing data values are quite common in air quality data. Such missing data values create complications in programming which must be studied in detail and assessed for possible corrections. It is essential that refinements or modifications be made in the computer program to compensate for missing data values.

Extensive analysis of the air quality data must be performed, especially with regard to detection of unrecognized periodic patterns and of relationships among pollution levels at different sites. This must be done in order to see what generalizations may be made. An attempt must be made to introduce weather factors (primarily temperature and wind velocity) into the analysis, in order to search for a relationship among these weather factors and levels of various pollutants.

The need for continuing refinements and modifications will become apparent as the actual analyses of production runs of the current computer program progress. And, even after the analyses of the results from the processed data have been largely completed, new insights will identify needs for further refinements and modifications of the then existing computer program packages.

V. REFERENCES

1. SAROAD Users Manual, Gerald Nehls, Environmental Protection Agency, Office of Air Programs, Research Triangle Park, N. C., July 1971, Publication Number APTD-0663.
2. SAROAD Station Coding Manual for Aerometric Sampling Networks, Donald H. Fair, Office of Air Programs, Research Triangle Park, N. C., February 1972, Publication Number APTD-0907.
3. VirginiaAir, quarterly publication of the Virginia State Air Pollution Control Board, Room 1106, Ninth Street Office Building, Richmond, Virginia 23219.
4. The Lognormal Distribution With Special Reference to Its Uses in Economics, J. Aitchison and J. A. C. Brown, New York: Cambridge University Press, 1957.
5. Introduction to Statistical Methods, 2nd. Ed., Donald L. Harnett, Reading, Massachusetts: Addison, Wesley, 1970, 1975.
6. Statistical Analysis for Decision Making, 2nd Ed., Morris Hamburg, New York: Harcourt, Brace and Jovanovich, 1970, 1977.
7. Statistics: Probability, Inference and Decision, William L. Hays and Robert L. Winkler, New York: Holt, Rinehart and Winston, 1971.
8. Statistical Analysis With Business and Economic Applications, 2nd. Ed., Ya-lun Chou, New York: Holt, Rinehart and Winston, 1969, 1975.
9. Data Processing of Air Pollution Data. written and oral report submitted at the termination of the 1976 NASA-ASEE LRC Summer Faculty Fellowship period by Dr. Charles M. Butler, Professor of Business Statistics and Data Processing, Mississippi State University, Mississippi State, Mississippi 39762, 8/13/76.
10. Processing Air Quality Data. unsolicited proposal submitted to the NASA Proposal Control Office by Dr. Charles M. Butler, Professor of Business Statistics and Data Processing, Mississippi State University, Mississippi State, Mississippi 39762, 8/18/76.
11. A Computer Program for Statistical Analysis of Environmental Data, Contract NAS1-14900, submitted to the Langley Research Center by T. Selden, Member of the Technical Staff, Computer Sciences Corporation, Hampton, Virginia 23666, revised 2/1/77.
12. Semi-Annual Status Report (January-June 1977) on NASA Grant NSG 1375: Processing Air Quality Data, submitted to Mr. John E. Hogge, NASA Technical Officer, Langley Research Center, by Dr. Charles M. Butler, Professor of Business Statistics and Data Processing, Mississippi State University, Mississippi State, Mississippi 39762, 7/29/77.

13. Processing Air Quality Data II, unsolicited proposal submitted to Mr. John E. Hogge, NASA Technical Officer, Langley Research Center, by Dr. Charles M. Butler, Professor, and Dr. Thomas E. Pritchard, Associate Professor of Business Statistics and Data Processing, Mississippi State University, Mississippi State, Mississippi 39762, 8/8/77.
14. Extension of the Multiplicative Time Series Model to Include Daily and Weekly Periodic Components, Working Paper Number 2, Charles M. Butler, Department of Business Statistics and Data Processing, Mississippi State University, Mississippi State, Mississippi 39762, 9/5/77.
15. Nonparametric Statistics for the Behavioral Sciences, Sidney Siegel, New York: McGraw-Hill Book Company, 1956.

ATTACHMENT

to

FINAL TECHNICAL REPORT
on
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
GRANT NUMBER NSG 1375
for
PROCESSING AIR QUALITY DATA

COMPUTER PROGRAM USER'S MANUAL

by

Dr. Thomas E. Pritchard, Associate Professor
Department of Business Statistics and Data Processing
Mississippi State University
Mississippi State, Mississippi 39762

May 22, 1978

Introduction

The programming package described in this documentation gives considerable flexibility to the user who wishes to analyze air pollution data. The only assumptions made about the data file are that it is in SAROAD format and that the data for any particular pollution parameter, area, and site are stored on the file in ascending chronological order. As presently constituted the program handles only "deblocked" files (one logical record/physical record), but this restriction could be removed by some minor modifications.

The program calculates a number of summary statistics, including the mean, variance, geometric mean, and coefficients of skewness and kurtosis. These statistics are calculated on a daily, monthly, or annual basis (average figures for time spans longer than one year may also be determined). Several types of frequency distributions are determined and goodness-of-fit tests of data to the normal, log-normal, gamma, and exponential distributions are made.

Also the program determines multiple regression and correlation coefficients, if it is desired to analyze relationships between several variables. Using the concept of moving averages, seasonal and weekly indices are also determined.

Finally, the program has limited plotting capabilities. One variable may be plotted against another, or a variable may be plotted against time.

The documentation consists of five sections:

- (1) POTENTIAL PROGRAM CHANGES BY THE USER
- (2) USER INPUT CARDS
- (3) SUBROUTINES
- (4) DETAILED DESCRIPTION OF PROGRAMS
- (5) APPENDIX

The first section describes a few changes that the user may easily make in the set of programs. The second section is the most immediately useful from the user's point of view. It gives a detailed description of the data cards the user needs to furnish for the purpose of indicating to the programs exactly what type of output he wants. Five examples are given of various user inputs, and the type of output that results from each input request is shown in the Appendix.

The third section gives a short description of the major purpose of each subroutine. The fourth section presents a more detailed description of each subroutine.

Before we conclude this introduction the following definition is in order. This definition will be used throughout the rest of the documentation. Period T: Period T will be used to refer to the period for which the user requests data analysis. For example, if summary statistics are requested for the period 4/1/75 to 7/1/77, period T would refer to the time span from 4/1/75 to 7/1/77.

Let X_{ij} = observation j on day i , where i is measured from the beginning of period T. (For example X_{14} would represent the 4th observation on the first day of period T).

$\bar{X}_{i.}$ = the daily mean for day i

$\bar{X}_{.j}$ = the mean of observation j for all of period T (that is
 $\bar{X}_{.j} = \sum_{i=1}^n X_{ij} / n$, where n is the total number of days in period t).

$\bar{X}_{.j}^1$ = the mean of observations j for one month.

Potential Program Changes by the User

There are two types of change which the user may wish to make in the program. Both types may be made by changing appropriate program cards in

the main routine. Thus these cards are in the first twenty cards of the deck.

The first type of change regards the logical unit numbers used for input units and output units in the program. The variable IDEV represents the number of the logical unit used to read in the air pollution data file. The statement:

IDEV=8

indicates that 8 is the number of the logical unit used. This statement may be changed to conform with the user's needs. The variable IN represents the number of the logical unit used to read in the user's input cards, which describe what printed reports are desired (see section USER INPUT CARDS for more detail about these cards). The statement:

IN=5

indicates that 5 is the number of the logical unit used. This statement may be changed to conform with the user's needs. The variable IOUT represents the number of the logical unit used to print out all desired information. The statement

IOUT=6

indicates that 6 is the number of the logical unit used. This statement may be changed to conform with the user's needs.

The second type of change regards the maximum sizes of certain arrays used in the programming package. The variable MAXSZ2 represents the maximum number of days for which data can be analyzed. The statement

MAXSZ2=1464

indicates that period T cannot be over 1464 days (4 years) in length. If a longer period than this is desired, the statement must be changed. The variable MAXSZ1 represents the maximum number of variables (independent + dependent) that can be used in the regression analysis. The statement

MAXSZ1=7

indicates that there cannot be more than 7 variables in the regression analysis. If more variables are desired, the statement must be changed. Also, if the value of MAXSZ1 or MAXSZ2 is changed the Dimension statements for the arrays DAØFMØ, MØØFYR, YEARTP, and DAXBAR need to be changed in the STAT1 routine. Likewise, the Dimension statements for the X, EX, and Y arrays need to be changed in the CORREL routine in this case.

User Input Cards

The user needs only a few data cards to indicate exactly what data he wants to have analyzed, and what type of output he wants. The first data card will indicate which routine should be called (STAT1 or CORREL), STAT1 giving most summary statistics, and CORREL giving correlation and regression information, and indices based on moving averages (see descriptions of these routines in the SUBROUTINES section of the documentation). Data cards after the first one will indicate the information that should be retrieved from the air pollution data file and specifically the output that should be generated. Both routines mentioned above can be called in the same run, so that there may be several sets of data cards of the type described below.

First card - Either 'SML' or 'REG' should be punched in columns 1-3, indicating calls to the STAT1 and CORREL routine, respectively.

I. If first card had 'SM1' in columns 1-3:

- A. Second card - this indicates what information is to be retrieved from the air pollution data file. Its setup is very similar to the setup of cards in SAROAD format, to minimize confusion. Following is the information on that card.

Col. 1 - a number that is 1 if the data are collected more frequently than every 24 hours; the number is 2 otherwise

Col. 2-3: state code

Col. 4-7: area code

Col. 8-10: site code

Col. 15-16: The number of the first year for which data are to be retrieved from the pollution file.

Col. 17-18: The number of the first month for which data are to be retrieved.

Col. 19-20: The number of the first day for which data are to be retrieved.

Col. 22-23: The number of the last year for which data are to be retrieved.

Col. 24-25: The number of the last month for which data are to be retrieved.

Col. 26-27: The number of the last day for which data are to be retrieved.

For example, if we wanted to retrieve data for the period from April 15, 1974 to June 28, 1975 we would punch, in columns 15-27: 740415 750628.

The period described by columns 15-27 will be referred to as period T (see definition in Introduction).

Col. 29-33: The code for the pollution parameter (for example, 42401 for sulfur dioxide).

- B. Third card - This card indicates what summary output is desired. If this card is left blank, all potential output from the STAT1 routine will be generated, which includes the following:
1. Listing of the information from the 80-column records from the air pollution data file (i.e., the "raw" data) for period T.
 2. Daily statistics, including the number of observations, daily mean, variance, coefficient of skewness, coefficient of kurtosis,

mean of the natural logs of the observations, log variance, geometric mean, minimum, maximum, range of observations, mid-range, and median observation.

3. Monthly statistics, which include the same statistics as the daily statistics, except that the average absolute deviation is calculated, and the median is not calculated.
4. A frequency distribution for the individual observations and a frequency distribution for the logs of the observations, for each month. For width of intervals, see writeup on FRQ routine. Also the chi-squared statistic, for attempted fits to the normal, log-normal, gamma, and exponential distributions, is printed.
5. i) A frequency distribution for the observations adjusted by subtracting from each observation the daily mean of the observation, (i.e., $X_{ij} - \bar{X}_{i.}$). Also a frequency distribution for the observations adjusted by subtracting $\bar{X}_{.j}$ from each observation is printed. The Chi-squared statistic for attempted fits of this adjusted data to the normal distribution is also printed. The frequency distributions calculated here are based on one month of data.
 ii) The means $\bar{X}_{.j}$ during the month (e.g., in the situation where there are 12 daily observations, the mean for all of the 2 a.m. observations for the month will be printed out, the mean for all of the 4 a.m. observations will be printed out, etc.)
6. The annual summary data. Exactly the same statistics as the monthly statistics, except for the average absolute deviation, are printed.
7. The same sorts of frequency distribution as in (4) above for all observations for an entire year.
8. The mean for each day of the week, these means being based on all observations over period T.
9. The same means as in (5ii), except that they represent all observations for period T instead of one month's observations (i.e., the means are the values $\bar{X}_{.j}$).
10. The frequency distributions for the daily means ($\bar{X}_{i.}$) and the logs of the daily means over all of period T. Also, the chi-squared statistics for attempted fits of these daily means to the normal, log-normal, gamma, and exponential distributions are printed.

11. A plot of daily means ($\bar{X}_t$) on the Y-axis, with time on the X-axis.

If the third card is not left blank, a punch of 1 in column 1 of the card will result in suppression of output (1) mentioned above; a punch of 1 in column 2 of the card will result in suppression of output (2) mentioned above, etc. For example a card that is blank except for a 1 in column 6 would result in all data being printed, except for the annual summary data mentioned in (6) above.

If a 2 is punched in column 3, the normal identification of area, site, pollutant, etc., that appears in the upper left-hand corner of the monthly summary page will be suppressed except for the first month. The user would normally only punch a 2 in column 3, when he was suppressing the outputs mentioned in (2), (4) and (5) above.

II. If first card had 'REG' in columns 1-3:

- A. Second card - In columns 1-2 the number of variables to be used (including the dependent variable) in the multiple regression and correlation should be punched. In column 5 a, "D" should be punched if the daily averages are to be treated as observations. If, however, the individual observations are to be used in calculating the regression coefficients and correlation coefficient a blank should be left in column 5.

If, however, the first variable represents time (see third paragraph below) the daily averages will be treated as individual observations, regardless of whether a "D" is punched in column 5.

On the same card, the user can indicate some of the output options he desires. Normally the output will include (1) a listing of the 80-column records from the air pollution data file for period T; (2) the multiple regression equation and multiple correlation coefficient; and (3) the plot of variable number two vs. variable number one (see below). If column 6 of card 2 contains '1', the 80-column listing is suppressed. If column 7 contains '1' the plot is suppressed. If variable number one represents time then the seasonal and weekly indices will be printed (if the total time period is at least 2 years) but the plot will be suppressed.

- B. There should be an additional card for each variable in the analysis. Thus, if 06 was punched in columns 1-2 of the second card, there should be 6 cards of the type described below. These cards will be referred to as "variable" cards. If time is a variable in the regression and correlation, the card corresponding to it should be first. The card corresponding to the dependent variable should be the last of the cards of the type described below.

The format of the card for each variable should be exactly the same as the format of the second card used by the STAT1 routine (described in (IA) above). If columns 29-33 are left blank in the first "variable" card, then time is treated as the first independent variable.

To recognize the flexibility in analysis that these "variable" cards allow, assume for a moment that there is one independent and one dependent variable. Thus there will be two "variable" cards. If these cards contain the same numbers, except for the site code in columns 8-10, then data for two different sites will be compared for the same pollutant and the same time period. If the pollutant code is the only number different on the two cards, then data for two different pollutants for the same site and time period will be correlated. If the time period is the only item different on the two cards, then data for the same pollutant and site will be compared for two different time periods.

We will now consider several examples of these data cards. We will explain the meaning of the numbers and will provide sample output that results from the requests by the data cards.

Example 1: SM1
 1480710001 721101 730430 42401
 1 1 1

The first card, with SM1 in columns 1-3, indicates that the report representing summary information for one variable is desired.

Second card: '1' in column 1 indicates data are collected with frequency greater than once every 24 hours.

'48' in columns 2-3 indicates Virginia.

'0710' in columns 4-7 indicates Chesapeake

'001' in columns 8-10 indicates Armory.

'721101' in columns 15-20 indicates that the beginning of the period for

which we want summary statistics calculated (i.e., period T) is Nov. 1, 1972.

'730430' in columns 22-27 indicates the end of the period is April 30, 1973.

'42401' in columns 29-33 indicates that we want summary statistics for sulfur dioxide for the period indicated.

Third card: The '1' in column 1 indicates that the listing of the raw data from the 80 column input records in the air pollution data file will be suppressed. That is, the raw data from the air pollution file for the period 11/01/72 to 04/30/73, on which the summary statistics will be based, will not be listed. The '1' in column 8 indicates that the mean for each day of the week, for the period 11/01/72 to 04/30/73, will not be printed. The '1' in column 11 indicates that the daily means for the period 11/01/72 to 04/30/73 will not be plotted. All other summary output for the one-variable case will be printed. See Section 1 in Appendix for the output generated.

Example 2: Now suppose that the first two cards are the same as in

Example 1, and the third card contains:

11111110110

Since there is a '1' in every column except for column 8 and 11, all output will be suppressed, except for the mean for each day of the week, and the plot of the daily means. See Section 2 in Appendix for the output generated.

Example 3: Let us consider a final example of the use of the STAT1 routine.

Assume that the third card contains:

01211000000

In this case the daily statistics will be suppressed, all frequency distributions for monthly statistics will be suppressed and item 5(ii) will be suppressed. See Section 3 in Appendix for the output generated.

Example 4: Now suppose that four data cards contain the following information:

```
REG
02 D
1480710001    721101 730430 42401
1480710001    721103 730502 42401
```

The first card has 'REG' in columns 1-3, which is a request for correlation/regression analysis. The '02' in columns 1-2 of the second data card indicates that there are to be two variables in the correlation and regression analysis. The 'D' in column 5 indicates that if the frequency of observations is greater than once per day, the daily averages will be calculated and used in place of the individual observations for the "raw data" in the correlation/regression. The information on the third card is exactly the same data as is on the second card in Example 1, and has exactly the same meaning. Note that the information on the fourth card is the same as is on the third card, except that the beginning and ending dates have been changed slightly (both have been lagged or shifted to the right two days). The effect is that the sulfur dioxide level for 11/01/72 will be compared with the level for 11/03/72, the level for 11/02/72 will be compared with the 11/04 level,..., the level for 4/30/73 will be compared with the level for 5/02/73. The third card represents variable number 1; the fourth card, variable number 2. The output that will result is indicated in Section 4 of the Appendix.

Example 5: Suppose that four data cards contain the following information:

```
REG
02 D
1480710001 721101 730430
1480710001 721101 730430 42401
```

The first two cards are the same as in Example 4. Note, however, that the beginning and ending dates on card 4 are different than the dates on card 4 in Example 4. These dates are the same as they are on card 3. Most importantly, note that columns 29-33 on card 3, which normally contain the pollution parameter code, have been left blank. The blank for the parameter code means that the first variable will represent time. The data for the second variable, which is represented by card number 4, will automatically be aggregated into daily averages, and the daily averages will be treated as individual observations in the regression of this variable against time.

The regression equation and multiple correlation coefficient will be printed, as in Example 4. Although ordinarily the seasonal and weekly indices will be printed when the first variable represents time, the calculation and printout of these indices are suppressed in this case. This is due to the fact that the total length of period T is less than 2 years. (Two years is the minimum period length required in order for seasonal and weekly indices to be calculated).

On both examples 4 and 5, columns 6 and 7 on the second card are left blank. The fact that column 6 is blank means that the data from the 80-column input records of the air pollution data file will be listed. The fact that column 7 is blank in example 4 means that a plot of the second variable (on the Y-axis) vs. the first variable will be made. However, since the first variable in example 5 represents time, no plot will be generated in example 5, even though column 7 is blank. (Output

option 11 of the summary statistics output allows for plotting of daily means of a pollution parameter, with time represented on the X-axis).

Important note: If the air pollution data is also read in via card file, then the user input cards described under (I) or (II) above should precede the air pollution data. If summary statistics calculations (as described under I) and regression analysis (as described under II) are to be done in the same run, then the cards requesting summary statistics should immediately precede the air pollution data to which the request is applicable. Likewise, the cards requesting regression should immediately precede the air pollution data to which the request is applicable.

Subroutines

1. Main program
2. STAT1
3. DAYS
4. FINDDA
5. CORREL
6. FRQ
7. GOOFIT
8. PLOT
9. GAMFIT
10. INV

Main routine - the main routine is a very short one that indicates what numbers are to represent the two input units and one output unit, the sizes of some of the major arrays, and what type of analysis is to be performed on the data. If the data card to indicate the type of analysis has 'SM1' in its 1st three columns the STAT1 routine will be called by the main program (see STAT1 description to indicate the function of the STAT1 routine). If the data card to indicate the type of analysis has 'COR' or 'REG' in its first three columns the CORREL routine will be called by the main program (see CORREL description). See the section "USER INPUT CARDS" to find out more about data card mentioned above.

STAT1 - The STAT1 routine is used to do most of the major summary calculations on the pollution data. It provides calculations of:

- A) daily summary data
- B) monthly summary data
- C) annual summary data

- D) frequency distributions for: individual observations; logs of individual observations; adjusted observations; daily means
- E) goodness of fit to normal, gamma, and exponential distributions for frequency distributions mentioned in (D).
- F) averages for the same time each day during period T.
- G) averages for each day of the week during period T.

It also will provide limited plotting capabilities. (See section "USER INPUT CARDS" and Examples 1-3 for more complete description of each of the items above).

The STAT1 routine calls the DAYS, FINDDA, FRQ, GOOFIT, PLOT, and GAMFIT routines. For further information about these, see the paragraphs on each of these routines.

DAYS - The DAYS routine receives a month number, day, and year through its argument list (variables IM, JD, and KY, respectively) and derives a value for a variable (L) which represents the days elapsed in the 20th century. This is so that consecutive days may be recognized as such when filling up arrays (for example 03/01/77 is recognized as being 1 day later than 2/28/77. Also days of the week (Monday, Tuesday, etc.) may be properly ascertained when necessary (see item G under description of STAT1 routine).

FINDDA - The FINDDA routine is referenced by both the STAT1 and CORREL routines. It is used to read the file of air pollution data, retrieving information from the file as requested in the STAT1 or CORREL routine.

CORREL - The CORREL routine is used to do any correlation/regression calculations requested (see write-up on main routine and the section "USER INPUT CARDS" to see how to request correlation/regression). If the independent variable in the regression is Time, then moving averages are also calculated on a "seasonal" basis (in reality an index is determined for each day of the year) and on a weekly basis (i.e., an index value is calculated for each day of the week). The CORREL routine calls the DAYS, FINDDA, INV and PLOT routines.

FRQ - The FRQ routine will take a given set of observations, along with their mean and variance, and construct a frequency distribution of these observations. As set up now, 18 intervals are used, the first interval representing all observations more than 2 standard deviations below the mean. The second through 17th intervals each have width $1/4S$ (e.g. the second interval represents all observations between 1.75 and 2 standard deviations below the mean, etc.) The 18th interval represents all observations greater than 2 standard deviations above the mean. Routine is called by the STAT1 routine.

- GOOFIT - The GOOFIT routine accepts the frequency distribution generated by the FRQ routine and calculates the chi-squared value for a test for goodness-of-fit of the observations to the normal distribution. A goodness-of-fit test for the logs of the observations is also done. Routine is called by the STAT1 routine.
- PLØT - The PLØT routine will generate plots (via the line printer) of one variable vs. another. Routine is called by STAT1 and CORREL.
- GAMFIT - The GAMFIT routine accepts the frequency distribution generated by the FRQ routine and calculates the chi-squared value for a test for goodness-of-fit of the observations to the Gamma distribution and to the exponential distribution. Numerical integration by Simpson's rule is used in order to evaluate probabilities for the Gamma distribution.
- INV - The INV routine handles matrix inversion. It is called by the CORREL routine to invert a matrix, in the process of finding the multiple regression coefficients.

DETAILED DESCRIPTION OF PROGRAMS

Note: The meanings of a few of the more important variables are indicated by comment cards inserted in the program deck. However, the general flow of the programs is indicated by the following section.

SUBROUTINE STAT1

Cards 710- This statement reads the state, area, site, pollution
720 parameter, and time period (period T) for which analysis
is requested.

Card 730 The IWRØPT array indicates which of the summary output will
be printed, and which will be suppressed.

Card 790 JTØTDY will be the total span, in days, of the period for
which data are to be analyzed (period T).

Card 900 Call FINDDA to read some records from air pollution data
file.

Cards 980- If NUMDAY=0, no data for period T was found, so error
990 message is printed, and execution is terminated.

Cards 1000- IOBS is the number of daily observations.
1040

Cards 1620- After all necessary data have been read for year, or data
3610 have been read for the last day of period T, most of the
statistics are calculated. Each pass through the DO loop
250 represents one month; each pass through loop 240
represents 1 day of the given month.

Cards 1320- Do loops 40, 44, and 48 determine area, site, and pollution
1470 parameter names that correspond to codes for the data that
are being analyzed.

Card 1790 NØDAYS is the number of days in given month.

Cards 1890- Each pass through Do-loop 230 corresponds to one observa-
1960 tion for the given day.

Cards 2090- Do loop 30 sorts the daily observations.
2160

Cards 2180- AMED, RANGE, and AMIDRA are median, range, and mid-range
2210 of daily observations, respectively.

Cards 2220- Calculate daily mean, variance, log mean, log variance,
2510 geometric mean, skewness, and kurtosis. Also add to running
totals for monthly statistics.

- Cards 2720- Do-loop 60 calculates the average absolute deviation from
2740 the mean for the month.
- Cards 2750- Calculate monthly summary statistics.
2900
- Cards 3100- Do loop 98 calculates the means $\bar{X}_j$ for the observations
3210 for the month (output item 5(ii) under Section I of
"USER INPUT CARDS").
- Cards 3260- Do loop 101 adjusts individual observations for the month
3460 by subtracting the daily means on one pass and by subtract-
ing the means $\bar{X}_j$ on the second pass.
- Cards 3620- Next 15 statements finish calculation of yearly summary
3760 statistics.
- Cards 3860- Do loop 145 calculates the frequency distribution of all
4050 observations for the year. Also goodness-of-fit tests
are done.
- Card 4210 Statements beyond this point are only executed after all
necessary data for period T have been read. Various
summary figures for the entire period are computed.
- Cards 4240- Statements down through statement 170 are to calculate and
4490 print out average pollution for each day of the week, for
all of period T (these averages are output item I, B(8) in
the description of "USER INPUT CARDS").
- Cards 4670- The frequency distributions for the daily means and the
4940 logs of the daily means, for all of period T are calculated.
Also goodness-of-fit tests are done. The printout for this
is mentioned in item I, B(10), under the description of
"USER INPUT CARDS".
- Cards 4960- A plot of the daily means vs. time (month, day, and year)
5050 is done for all of period T.

SUBROUTINE FINDDA

- Cards 5770- These two read statements are used to read records from the
5820 air pollution data file. The first read statement is used
for records which contain a "1" in column 1 (frequency of
observations greater than once/day). The second read
statement is used for records that contain "2" in column 1.
- Card 5840 If the CORREL routine called the FINDDA routine, the rest
of FINDDA is skipped. The checking (to see if the record
just read corresponds to information requested) is done in
the CORREL routine.

Cards 5880- If SAROAD data contains "2" in column 1, there may be data
5980 for up to 4 pollutants on same record. Record just read is checked to see if pollutant for which analysis was requested matches any of the (up to 4) pollutants mentioned on SAROAD record.

Card 6000 This point is only reached if pollution parameter, site, and area on SAROAD record match those for which analysis was requested. If the date on the SAROAD record also is contained in period T, the RETURN statement will be executed, and data from the SAROAD record will be passed back to the calling routine.

SUBROUTINE CORREL

Card 6430 IEND(I) is 0 until pollution data for variable I have been read from the air pollution file for a date beyond the last day of period T for variable I. When IEND(I) is non-zero for all I, there is no further need to read the file.

Card 6500 The X(I,J) array contains the various observations for each variable (I) over the time span of period T.

Card 6540 NOVAR is the total number of variables. AGGREG is the aggregation level for the data (if AGGREG is 'D' all observations for one day will be averaged together and the daily average will be treated as one observation). The array IWROPT will control the output that will be printed (see writeup USER INPUT CARDS).

Cards 6610- The Read statement reads the parameters that define each
6620 variable in the regression problem. Submission of a card that is blank (except for the indicated beginning and ending of the period for which data is to be analyzed) for the first variable will cause the first variable to be treated as time.

Card 6700 If first variable represents time, data for dependent variable will be aggregated on daily basis (daily averages will be treated as individual observations).

Cards 6720- A check is made to see if the arrays have been dimensioned
6740 large enough to allow for the time span (in days) of period T.

Cards 6790- This Do-loop checks to see if the record just read by FINDDA
7370 corresponds to a parameter requested, in the time period requested. It also fills in the X array. If the number of individual observations for each variable over the time period for which data are to be analyzed exceed the value of MAXSZ2, all observations on one day are averaged (the daily average then being treated as one observation).

Cards 7120- The nine statements aggregate the data for one day into one
7200 daily average, which is then treated as one observation.

- Cards 7220- This Do-loop is executed if the data are not to be aggregated
7310 into daily averages, or if the frequency of observations is one/day or less.
- Cards 7380- This Do-loop checks to see whether the most recent record
7400 read from the tape corresponds to a date that is after the last date for which analysis is requested for any variable.
- Cards 7420- This Do-loop fills up the first row in the X array
7430 (representing the first variable's values) when the first variable is time. One unit represents 30 days. Since $X(1, I+1)$ represents one day later than $X(1, I)$ the difference between $X(1, I+1)$ and $X(1, I)$ is $1/30$.
- Cards 7470- These Do-loops calculate the sum of the observations for
7570 each variable and the sum of the cross-products of the observations. If an observation is missing for one variable, the corresponding observation for the other variables is discarded.
- Cards 7720- The routine INV is called to invert the matrix of sums and
7790 cross-products in order to determine the coefficients of the regression equation and the multiple correlation coefficient. The coefficients of the regression equation are returned as elements of column NOVAR+1 of the A array. The variable R represents the multiple correlation coefficient.
- Cards 7830- Do-loops 81, 84, and 88 determine area, site, and pollution
7970 parameter names that correspond to the codes for the data that are being analyzed.
- Cards 8000- The set of Write statements in Do-loop 70 and immediately
8110 following that write out a verbal identification for each variable and then the regression equation and correlation coefficient.
- Cards 8120- If the first variable does not represent time or the length
8130 of period T is less than two years, the seasonal and cyclical indices are not calculated (the "GO TO 120" is executed).
- Cards 8140- Before doing any time series analysis, data gaps are
8350 "filled in" by the procedures in these statements.
- Cards 8360- This group of statements calculates an index number for
8440 every day, beginning with the 183rd day of period T and ending with the 182nd day prior to the end of period T (i.e., MAXPOS-182, where MAXPOS is the total number of days in the period). The index is calculated by dividing the daily mean by the moving average centered on that day.
- Cards 8470- These statements will check to see whether any day of the
8640 year has had more than one index calculated above. If any day has had two or more index numbers calculated for it (which will be the case if the total time span for which analysis is requested exceeds 2 years) the index numbers will be averaged to derive one index for that day.

- Cards 8650- The statements down through statement 300 will calculate
8840 weekly indices in the same general manner that seasonal indices were calculated.
- Cards 8860- Do-loop 340 will print out the 7 weekly index numbers,
8900 identified by day of the week.
- Cards 8920- Do-loop 330 will print out the 365 "seasonal" index
9050 figures, which will be identified by the number of the day of the year.
- Cards 9070- The rest of this subroutine is for the purpose of
9290 plotting values of variable 2 against values of variable 1, if such a plot is requested by the user's input cards. Do loop 350 sorts the pairs of observations so that the values of variable 1 are in ascending order. Variable 1 will represent the X-axis.

SUBROUTINE FRQ

- Cards 9740- This routine calculates the frequency distribution for
9750 observations and logs of observations. If ILOG is 0, the frequency distribution is based on the observations. If ILOG is 1, the frequency distribution is for the logs of the observations.
- Cards 9810- Do-loop 80 assigns each observation (or log of observation)
9870 to the appropriate interval (i.e. one of the class counts is increased by 1). There are 18 classes, or intervals, the width of all but the first and last interval being $1/4$ standard deviation. The first and last intervals represent all observations more than 2 standard deviations below, and more than 2 standard deviations above, the mean, respectively.
- Cards 9910- Do-loop 96 prints out the lower and upper bound for each
9950 interval, and the number of observations falling in each interval.

SUBROUTINE GOOFIT

- Cards 10150- Do-loop 20 computes the Chi-square statistic for goodness-
10250 of-fit of the data to a normal distribution. For each of the 18 intervals or classes, the expected Frequency (EXPFR) of observations in that class is compared with the actual frequency of observations (FREQ) in that class. If either the actual or expected frequency is less than 5 for any class, however, that class is combined with the next class. These two classes are treated as one, for the purpose of calculating the Chi-squared statistic. The number of degrees of freedom is represented by the variable IDF.

SUBROUTINE PLOT

- Cards 10560- Do-loop 20 computes 10 values of the YM array, which will
 10570 be printed along the Y-axis to identify the scaling on the
 Y-axis. The Y-axis will represent the values of variable
 number 2.
- Cards 10610- The seven statements here regulate the number of lines the
 10670 printer carriage should be moved down before the 1st point
 (representing (X_1, Y_1)) is plotted.
- Cards 10680- If the 1st value of variable number 2 (that is, Y_1) is
 10710 recorded (i.e., Y_1 is not 9999.) then the 1st point will be
 plotted by statement 70. Otherwise nothing will be plotted.
- Cards 10720- Do-loop 90 will take care of the plotting of all points
 10910 after X_1, Y_1 .
- Cards 10730- These four statements determine the number of lines to ad-
 10760 vance the printer carriage before plotting the next point.
 The total number of lines that the carriage has been
 advanced is also updated.
- Cards 10780- These five statements move the printer carriage down a line
 10820 at a time until it has been moved down a distance equal to
 the value of MOVE (which represents the number of lines down
 the X-axis between the values of X_i and X_{i+1}).
- Cards 10840- One of these two write statements will be executed to print
 10850 out values of X_1 (the 1st variable) to identify the scaling
 on the X-axis. If the 1st variable is time, the month, day,
 and year are printed for the scale values on the X-axis.
- Card 10900 If the i th value of variable number 2 is recorded (i.e., is
 not 9999.) then point i will be plotted by statement 150.

SUBROUTINE GAMFIT

- Cards 11090- The statements down through statement 20 are to calculate
 11610 the Chi-squared statistic for a fit to the Gamma distribu-
 tion. Eighteen classes (intervals) are used.
- Cards 11140- These 3 statements calculate the estimated parameters B and
 11160 C of the Gamma distribution, as functions of the mean and
 variance.
- Cards 11170- The statements down through statement 15 calculate the
 11280 constant part of the Gamma density function
 (i.e., $C^{b+1}/\Gamma(b+1)$).

- Cards 11330- One pass is made through Do-loop 20 for each of the classes
11610 or intervals. The expected frequency of observations in each interval is computed and compared with the actual number of observations in each interval, in order to calculate the Chi-squared statistic. If either the actual or expected frequency is less than 5 for any class, however, that class is combined with the next class. These two classes are treated as one, for the purpose of calculating the Chi-squared statistic. The number of degrees of freedom is represented by the variable IDF.
- Cards 11370- The expected number of observations in each class is
11490 evaluated by first finding the probability of an observation falling in that class. In order to find this probability each interval or class is divided into ten subintervals. Then Simpson's rule is used to numerically integrate the Gamma density function over the interval.
- Cards 11650- The rest of this subroutine is used to calculate the Chi-
11870 squared statistic for a fit to the exponential distribution. Eighteen intervals are used and the procedure is basically the same as for the Gamma distribution. However, numerical integration is not necessary, since a closed-form integral exists for the exponential density function. Thus, there is no need to divide each interval or class into 10 sub-intervals.

Example 1

AIR POLLUTION ANALYSIS
 STATE: 40 VA.
 AREA: 0710 CHESAPEAKE
 SITE: 001 ANAHDY
 POLLUTANT: 42401 SULFUR DIOX.
 MONTH AND YEAR: 11 72

DATE	TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MID-RANGE	AV. AB. DEV
11 1	23	.348-02	.227-04	.039	-1.53	-12.1	30.1	.588-05	.100-06	.100-01	.100-01	.500-02	.08
11 2	24	.317-01	.105-02	.994	-1.02	-16.62	30.5	.131-02	.100-06	.100-01	.100-01	.550-01	.00
11 3	24	.158-01	.020-04	2.25	5.67	-4.26	.200	.141-01	.100-01	.500-01	.400-01	.300-01	.00
11 4	24	.121-01	.100-03	.910	.410	-7.22	26.6	.733-03	.100-06	.400-01	.400-01	.200-01	.00
11 5	24	.187-01	.130-03	.075	-.867	-4.15	.324	.158-01	.100-01	.400-01	.300-01	.250-01	.00
11 6	23	.104-01	.324-04	-.557-01	-1.78	-8.81	34.4	.140-03	.100-06	.200-01	.200-01	.100-01	.00
11 7	24	.146-01	.332-04	.018	-.332	-4.30	.137	.136-01	.100-01	.500-01	.200-01	.200-01	.00
11 8	24	.167-01	.639-04	1.16	.024	-4.20	.193	.151-01	.100-01	.400-01	.300-01	.250-01	.00
11 9	24	.380-01	.200-03	-.354	-1.000	-3.63	.270	.265-01	.100-01	.500-01	.500-01	.350-01	.00
11 10	24	.254-01	.250-03	-.196	-1.35	-10.23	17.3	.260-02	.100-06	.500-01	.500-01	.350-01	.00
11 11	24	.750-02	.604-04	.998	.849	-9.30	37.3	.915-04	.100-06	.500-01	.500-01	.150-01	.00
11 12	24	.702-02	.330-04	.417-01	-.332	-7.91	27.8	.360-03	.100-06	.200-01	.200-01	.100-01	.00
11 13	24	.100-01	.117-03	1.58	5.26	-8.24	31.0	.250-03	.100-06	.500-01	.500-01	.250-01	.00
11 14	24	.150-01	.150-03	1.58	1.94	-10.22	11.1	.530-02	.100-06	.500-01	.500-01	.350-01	.00
11 15	24	.750-02	.854-04	1.15	.420	-10.2	35.0	.367-04	.100-06	.500-01	.500-01	.150-01	.00
11 16	24	.438-02	.490-04	1.50	-.200	-12.4	20.0	.399-05	.100-06	.200-01	.200-01	.100-01	.00
11 17	24	.456-02	.352-04	.338	-.332	-11.3	32.6	.125-04	.100-06	.200-01	.200-01	.100-01	.00
11 18	24	.100-01	.167-04	-.330	3.00	-10.3	30.3	.400-02	.100-06	.200-01	.200-01	.100-01	.00
11 19	24	.583-02	.243-04	-.022	-1.89	-9.40	32.2	.820-04	.100-06	.100-01	.100-01	.500-02	.00
11 20	24	.150-01	.433-04	-.022	-.636	-10.21	11.0	.584-02	.100-06	.400-01	.400-01	.200-01	.00
11 21	24	.240-01	.010-04	-.044	-.974	-10.82	10.3	.220-01	.100-01	.400-01	.300-01	.250-01	.00
11 22	24	.190-01	.120-03	1.70	-.019	-10.00	.105	.167-01	.100-01	.400-01	.300-01	.250-01	.00
11 23	24	.204-01	.100-03	-.124	1.37	-3.64	.105	.262-01	.100-01	.400-01	.300-01	.250-01	.00
11 24	24	.179-01	.740-04	1.19	.982	-4.13	.201	.181-01	.100-01	.400-01	.300-01	.250-01	.00
11 25	24	.154-01	.400-03	2.04	7.23	-7.64	30.0	.481-03	.100-06	.400-01	.100-01	.100-01	.00
11 26	24	.142-01	.224-03	2.32	4.47	-10.82	15.8	.297-02	.100-06	.500-01	.600-01	.400-01	.00
11 27	24	.148-01	.400-03	1.66	2.43	-8.58	34.4	.188-03	.100-06	.700-01	.700-01	.500-01	.00
11 28	24	.530-01	.130-02	1.53	-.961	-3.13	.784	.436-01	.100-01	.100-01	.130	.750-01	.00
11 29	24	.750-02	.521-04	.410	-1.01	-10.20	33.4	.924-04	.100-06	.500-01	.200-01	.100-01	.00
11 30	24	.503-02	.750-04	1.54	1.11	-12.6	20.6	.328-04	.100-06	.500-01	.300-01	.150-01	.00
MONTH	TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MID-RANGE	AV. AB. DEV
11 72	710	.159-01	.310-03	2.35	9.32	-6.86	25.8	.105-02	.100-06	.140	.140	.780-01	.01

ORIGINAL PAGE IS
 OF POOR QUALITY

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOX.

MONTH AND YEAR: 11 72

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-1950-01	0
-1950-01	-1512-01	0
-1512-01	-1087-01	0
-1087-01	-650-02	0
-650-02	-1814-02	0
-1814-02	2022-02	163
2022-02	7058-02	0
7058-02	1149-01	249
1149-01	1593-01	0
1593-01	2037-01	133
2037-01	2480-01	0
2480-01	2924-01	0
2924-01	3367-01	43
3367-01	3811-01	0
3811-01	4255-01	46
4255-01	4698-01	0
4698-01	5142-01	14
5142-01	9999.	24

CHI-SQUARED VALUE WITH 4 D. F. IS 330.97

TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS B= .01 AND C= 50.60

CHI-SQUARED VALUE WITH 4 D. F. IS 179.14

TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .02

CHI-SQUARED VALUE WITH 5 D. F. IS 80.57

FREQUENCY DISTRIBUTION FOR LOGS OF OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-17.02	0
-17.02	-15.75	163
-15.75	-14.40	0
-14.40	-13.21	0
-13.21	-11.94	0
-11.94	-10.67	0
-10.67	-9.404	0
-9.404	-8.133	0
-8.133	-6.863	0
-6.863	-5.593	0
-5.593	-4.323	249
-4.323	-3.053	220
-3.053	-1.782	38
-1.782	-0.5122	0
-0.5122	0.7580	0
0.7580	2.028	0
2.028	3.298	0
3.298	9999.	0

CHI-SQUARED VALUE WITH 10 D. F. IS 1155.40

ORIGINAL PAGE IS
OF POOR QUALITY

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 07111 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOX.

MONTH AND YEAR: 11 72

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
9999.	1750-01	0
1999-01	1512-01	0
1512-01	1069-01	0
1069-01	5250-02	0
6250-02	1814-02	0
1814-02	2022-02	103
2022-02	7050-02	0
7050-02	1149-01	249
1149-01	1593-01	0
1593-01	2037-01	133
2037-01	2484-01	0
2484-01	2924-01	0
2924-01	3367-01	43
3367-01	3811-01	0
3811-01	4255-01	44
4255-01	4698-01	0
4698-01	5142-01	14
5142-01	9999.	24

CHI-SQUARED VALUE WITH 40 D.F. IS 300.97

TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS $\alpha = .01$ AND $\beta = 50.60$

CHI-SQUARED VALUE WITH 4 D.F. IS 179.14

TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .02

CHI-SQUARED VALUE WITH 5 D.F. IS 80.57

FREQUENCY DISTRIBUTION FOR LOGS OF OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
9999.	17.02	0
17.02	15.75	163
15.75	14.40	0
14.40	13.21	0
13.21	11.94	0
11.94	10.67	0
10.67	9.404	0
9.404	8.133	0
8.133	6.863	0
6.863	5.593	0
5.593	4.323	240
4.323	3.053	220
3.053	1.784	38
1.784	0.512	0
0.512	0.240	0
0.240	0.020	0
0.020	9999.	0

CHI-SQUARED VALUE WITH 10 D.F. IS 1153.40

ORIGINAL PAGE IS
OF POOR QUALITY

AIR POLLUTION ANALYSIS

STATE: 46 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 4201 SULFUR DIOX.

MONTH AND YEAR: 11 72

OBSERVATIONS ADJUSTED BY SUBTRACTING DAILY AVERAGES

FREQUENCY DISTRIBUTION FOR OBSERVATIONS		
LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999	-2770-01	13
-2770-01	-2424-01	3
-2424-01	-2077-01	3
-2077-01	-1731-01	8
-1731-01	-1385-01	34
-1385-01	-1039-01	21
-1039-01	-6925-02	104
-6925-02	-3462-02	177
-3462-02	-6076-09	46
-6076-09	-3462-02	78
-3462-02	-6925-02	86
-6925-02	-1039-01	28
-1039-01	-1385-01	32
-1385-01	-1731-01	24
-1731-01	-2077-01	11
-2077-01	-2424-01	13
-2424-01	-2770-01	5
-2770-01	9999	24

CHI-SQUARED VALUE WITH 130 D.F. IS 355.27

OBSERVATIONS ADJUSTED BY SUBTRACTING 1-HOURLY AVERAGES

FREQUENCY DISTRIBUTION FOR OBSERVATIONS		
LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999	-3510-01	0
-3510-01	-3077-01	0
-3077-01	-2637-01	0
-2637-01	-2198-01	6
-2198-01	-1758-01	22
-1758-01	-1319-01	128
-1319-01	-8791-02	32
-8791-02	-4390-02	185
-4390-02	-1019-08	95
-1019-08	-4390-02	59
-4390-02	-8791-02	70
-8791-02	-1319-01	11
-1319-01	-1758-01	26
-1758-01	-2198-01	4
-2198-01	-2637-01	33
-2637-01	-3077-01	10
-3077-01	-3510-01	10
-3510-01	9999	25

CHI-SQUARED VALUE WITH 100 D.F. IS 513.14

AVERAGES FOR THE SAME TIME EACH DAY DURING THE MONTH

NUMBER 1	NUMBER 2	NUMBER 3	NUMBER 4	NUMBER 5	NUMBER 6	NUMBER 7	NUMBER 8	NUMBER 9	NUMBER 10	NUMBER 11	NUMBER 12
.0163	.0140	.0114	.0102	.0166	.0152	.0134	.0131	.0155	.0217	.0210	.0150
.0177	.0180	.0167	.0140	.0170	.0163	.0143	.0173	.0147	.0140	.0133	.0150

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOX.

MONTH AND YEAR: 12 72

DATE	TALLY	MEAN	VARIANCE	COEFF. SKENNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MTD-RANGE	AV. AB. DEV
12 1	23	.696-02	.907-04	1.05	3.75	-10.5	34.5	.278-04	.100-06	.400-01	.400-01	.200-01	.00
12 2	24	.103-01	.490-03	2.11	3.67	-7.02	30.1	.491-03	.100-06	.400-01	.400-01	.450-01	.01
12 3	24	.504-01	.044-02	3.27	10.5	-4.38	7.28	.130-01	.100-06	.460	.460	.230	.01
12 4	24	.179-01	.700-03	2.52	5.48	-7.13	27.4	.808-03	.100-06	.110	.110	.550-01	.01
12 5	24	.400-01	.124-02	1.74	-1.753	-4.08	7.14	.169-01	.100-06	.110	.110	.550-01	.03
12 6	24	.117-01	.289-03	1.03	3.35	-10.5	37.5	.277-04	.100-06	.700-01	.700-01	.350-01	.00
12 7	24	.833-03	.764-05	3.02	7.09	-15.2	10.1	.261-06	.100-06	.100-01	.100-01	.500-02	.00
12 8	24	.100-01	.267-03	1.05	1.80	-11.5	35.6	.101-04	.100-06	.600-01	.600-01	.300-01	.00
12 9	24	.700-02	.707-04	1.43	1.84	-9.75	34.0	.560-04	.100-06	.300-01	.300-01	.150-01	.01
12 10	24	.125-02	.109-04	2.27	3.14	-14.7	14.5	.422-06	.100-06	.100-01	.100-01	.500-02	.00
12 11	24	.750-02	.854-04	1.78	3.85	-9.77	34.2	.578-04	.100-06	.400-01	.400-01	.200-01	.01
12 12	24	.273-01	.874-03	1.38	-1.73	-6.57	30.9	.190-03	.100-06	.700-01	.700-01	.350-01	.01
12 13	24	.354-01	.825-03	.738	-1.383	-5.04	19.0	.646-02	.100-06	.100-00	.100-00	.500-01	.02
12 14	24	.100-01	.106-02	-1.66-12	.000	-10.1	-1.763-05	.100-06	.100-06	.100-06	.000	.100-06	.00
12 15	24	.417-02	.493-04	1.38	.399	-12.7	28.8	.313-05	.100-06	.200-01	.200-01	.100-01	.00
12 16	24	.417-03	.399-05	4.59	-19.0	-15.6	5.29	.162-06	.100-06	.100-01	.100-01	.500-02	.00
12 17	24	.100-01	.142-03	.741	-1.027	-10.1	36.6	.421-04	.100-06	.300-01	.300-01	.150-01	.01
12 18	24	.910-02	.601-04	1.67	.007	-7.51	26.2	.546-03	.100-06	.300-01	.300-01	.150-01	.01
12 19	24	.250-01	.567-03	2.90	8.26	-3.93	38.1	.196-01	.100-01	.120	.110	.650-01	.02
12 20	24	.504-01	.240-02	1.22	-1.475-01	-3.41	30.4	.230-01	.100-01	.170	.160	.200-01	.02
12 21	24	.375-02	.234-04	.915	-1.73	-11.4	31.1	.750-05	.100-06	.100-01	.100-01	.500-02	.00
12 22	24	.122-01	.104-03	.766	.439	-11.32	27.5	.642-03	.100-06	.400-01	.400-01	.200-01	.01
12 23	24	.375-02	.401-04	1.46	.898	-12.7	28.4	.344-05	.100-06	.200-01	.200-01	.100-01	.00
12 24	24	.625-02	.132-03	2.09	3.37	-12.1	31.8	.536-05	.100-06	.400-01	.400-01	.200-01	.00
12 25	24	.700-02	.140-03	2.14	4.39	-11.6	33.7	.889-05	.100-06	.500-01	.500-01	.250-01	.02
12 26	24	.250-01	.445-03	1.30	1.98	-5.08	14.8	.623-02	.100-06	.900-01	.900-01	.450-01	.02
12 27	24	.212-01	.175-03	1.46	1.84	-4.02	31.4	.170-01	.100-01	.600-01	.600-01	.350-01	.02
12 28	24	.158-01	.320-04	2.23	5.67	-4.26	20.0	.141-01	.100-01	.500-01	.500-01	.300-01	.01
12 29	24	.917-02	.320-04	-0.19-02	-1.271-02	-6.92	22.3	.941-03	.100-06	.200-01	.200-01	.100-01	.01
12 30	24	.625-02	.310-04	.196	-1.788	-9.37	32.5	.850-04	.100-06	.200-01	.200-01	.100-01	.01
12 31	24	.108-01	.160-04	4.59	19.0	-4.56	48.2-01	.105-01	.100-01	.300-01	.200-01	.200-01	.01
MONTH TALLY		MEAN	VARIANCE	COEFF. SKENNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MTD-RANGE	AV. AB. DEV
12 72	727	.145-01	.803-03	7.16	80.2	-9.04	35.5	.119-03	.100-06	.460	.460	.230	.01

AIR POLLUTION ANALYSIS
 STATE: 46 VA.
 AREA: 0710 CHESAPEAKE
 SITE: 001 ARROYO
 POLLUTANT: 42401 SULFUR DIOX.
 MONTH AND YEAR: 12 72

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999	-4216-01	0
-4216-01	-3500-01	0
-3500-01	-2799-01	0
-2799-01	-2091-01	0
-2091-01	-1383-01	0
-1383-01	-6742-02	0
-6742-02	3429-03	300
3429-03	7427-02	0
7427-02	1451-01	242
1451-01	2180-01	78
2180-01	2480-01	0
2480-01	3571-01	30
3571-01	4285-01	21
4285-01	4993-01	0
4993-01	5702-01	15
5702-01	6410-01	11
6410-01	7119-01	8
7119-01	9999	22

CHI-SQUARED VALUE WITH 60 D.F. IS 209.64

TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS $\alpha = .01$ AND $\beta = 18.07$

CHI-SQUARED VALUE WITH 5 D.F. IS 1180.36

TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .01

CHI-SQUARED VALUE WITH 5 D.F. IS 4927.32

FREQUENCY DISTRIBUTION FOR LOGS OF OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999	-20.89	0
-20.89	-19.47	0
-19.47	-17.90	0
-17.90	-16.42	0
-16.42	-15.00	300
-15.00	-13.51	0
-13.51	-12.02	0
-12.02	-10.53	0
-10.53	-9.04	0
-9.04	-7.55	0
-7.55	-6.06	0
-6.06	-4.57	242
-4.57	-3.08	126
-3.08	-1.59	54
-1.59	-.9999-01	2
-.9999-01	1.390	0
1.390	2.000	0
2.000	9999	0

CHI-SQUARED VALUE WITH 10 D.F. IS 543.99

ORIGINAL PAGE IS
 OF POOR QUALITY

AIR POLLUTION ANALYSIS
 STATE: 48 VA.
 AREA: 0710 CHESAPEAKE
 SITE: 001 ARMORY
 POLLUTANT: 42401 SULFUR DIOX.
 MONTH AND YEAR: 12 72

OBSERVATIONS ADJUSTED BY SUBTRACTING DAILY AVERAGES

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-4959-01	1
-4959-01	-4959-01	10
-4959-01	-3719-01	19
-3719-01	-3099-01	13
-3099-01	-2480-01	34
-2480-01	-1860-01	4
-1860-01	-1240-01	35
-1240-01	-0199-02	107
-0199-02	0501-09	194
0501-09	0899-02	128
0899-02	1240-01	50
1240-01	1860-01	26
1860-01	2480-01	17
2480-01	3099-01	8
3099-01	3719-01	11
3719-01	4339-01	5
4339-01	4959-01	1
4959-01	9999.	24

CHI-SQUARED VALUE WITH 100 D.F. IS 564.85

OBSERVATIONS ADJUSTED BY SUBTRACTING 1-HOURLY AVERAGES

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-5000-01	0
-5000-01	-4900-01	0
-4900-01	-4200-01	0
-4200-01	-3501-01	0
-3501-01	-2801-01	9
-2801-01	-2101-01	107
-2101-01	-1401-01	107
-1401-01	-7003-02	188
-7003-02	7003-02	06
7003-02	1401-01	37
1401-01	2101-01	17
2101-01	2801-01	18
2801-01	3501-01	12
3501-01	4200-01	0
4200-01	4900-01	4
4900-01	5000-01	26
5000-01	9999.	

CHI-SQUARED VALUE WITH 100 D.F. IS 755.12

AVERAGES FOR THE SAME TIME EACH DAY DURING THE MONTH

OBSERV NUMBER	1	2	3	4	5	6	7	8	9	10	11	12
0163	0104	0147	0143	0147	0157	0150	0155	0116	0081	0100	0077	
0163	0104	0147	0143	0147	0157	0150	0155	0116	0081	0100	0077	
0120	0134	0131	0105	0152	0177	0219	0283	0107	0117	0168	0173	

AIR POLLUTION ANALYSIS

STATE: 46 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOXIDE

ANNUAL SUMMARY: 1972

TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MID-RANGE
1437	.152-01	.562-03	6.67	93.4	-7.96	31.9	.348-03	.100-06	.460	.460	.230

ORIGINAL PAGE IS
OF POOR QUALITY

AIR POLLUTION ANALYSIS
 STATE: 45 VA.
 AREA: 0710 CHESAPEAKE
 SITE: 001 ARMORY
 POLLUTANT: 42901 SULF, RILIX.
 ANNUAL SUMMARY: 1972

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-3221-01	0
-3221-01	-2029-01	0
-2029-01	-2030-01	0
-2030-01	-1443-01	0
-1443-01	-3501-02	0
-3501-02	-2575-02	0
-2575-02	3350-02	403
3350-02	9284-02	0
9284-02	1521-01	531
1521-01	2114-01	211
2114-01	2707-01	70
2707-01	3300-01	73
3300-01	3894-01	0
3894-01	4485-01	65
4485-01	5070-01	29
5070-01	5671-01	0
5671-01	6264-01	14
6264-01	9999.	51

CHI-SQUARED VALUE WITH 50 D.F. IS 528.18

TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS B= .01 AND C= 27.05

CHI-SQUARED VALUE WITH 50 D.F. IS 1576.47

TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .02

CHI-SQUARED VALUE WITH 50 D.F. IS 257.37

FREQUENCY DISTRIBUTION FOR LOSS OF OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-19.20	0
-19.20	-17.83	0
-17.83	-10.44	0
-10.44	-13.63	403
-13.63	-13.61	0
-13.61	-12.20	0
-12.20	-10.73	0
-10.73	-9.377	0
-9.377	-7.964	0
-7.964	-6.552	0
-6.552	-5.140	0
-5.140	-3.726	742
-3.726	-2.310	104
-2.310	-9633	23
-9633	5082	1
5082	1.921	0
1.921	3.333	0
3.333	9999.	0

CHI-SQUARED VALUE WITH 1 D.F. IS 999.99

ORIGINAL PAGE IS
 OF POOR QUALITY

AIR POLLUTION ANALYSIS
 STATE: 46 VA.
 AREA: 0710 CHESAPEAKE
 SITE: 001 ARMORY
 POLLUTANT: 42401 SULFUR DIOX.
 MONTH AND YEAR: 1 73

DATE	TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MID-RANGE	AV. AD. DEV
1 1	24	.503-01	.724-02	4.27	17.6	-3.42	.500	.327-01	.100-01	.450	.440	.230	.02
1 2	25	.153-01	.405-04	.163	-6.17	-4.07	.143	.171-01	.100-01	.500-01	.200-01	.200-01	.02
1 3	25	.178-01	.131-03	.488-61	-9.25	-3.98	16.9	.378-02	.100-06	.500-01	.400-01	.200-01	.02
1 4	25	.240-01	.204-03	1.44	1.50	-3.98	.320	.131-01	.100-01	.500-01	.500-01	.350-01	.01
1 5	25	.183-01	.553-03	1.07	1.88	-7.68	.316	.402-03	.100-06	.900-01	.400-01	.450-01	.01
1 6	24	.125-01	.437-04	-3.4	-7.76	-5.78	15.8	.308-02	.100-06	.200-01	.200-01	.100-01	.01
1 7	24	.121-01	.248-04	2.38	-4.79	-4.47	.928-01	.114-01	.100-01	.500-01	.200-01	.200-01	.01
1 8	24	.107-01	.226-04	-7.07	-1.50	-4.14	.107	.159-01	.100-01	.200-01	.100-01	.150-01	.02
1 9	24	.321-01	.606-04	-0.76	-1.48	-3.51	.166	.300-01	.100-01	.500-01	.400-01	.300-01	.03
1 10	24	.308-01	.208-03	1.11	.570	-3.58	.196	.278-01	.100-01	.700-01	.600-01	.400-01	.01
1 11	24	.238-01	.107-03	.110	.602	-4.20	5.24	.137-01	.100-06	.500-01	.500-01	.250-01	.02
1 12	24	.101-01	.051-04	.303	.728	-4.07	5.88	.937-02	.100-06	.300-01	.300-01	.150-01	.02
1 13	24	.204-01	.537-03	1.05	1.16	-4.21	.501	.140-01	.100-01	.700-01	.600-01	.400-01	.01
1 14	24	.321-01	.175-02	1.51	1.91	-3.27	.605	.382-01	.100-01	.100	.170	.050-01	.02
1 15	24	.321-01	.311-03	1.83	3.17	-3.87	.356	.281-01	.100-01	.900-01	.400-01	.500-01	.04
1 16	24	.258-01	.241-03	1.70	1.81	-2.93	.351	.532-01	.300-01	.100	.170	.050-01	.02
1 17	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 18	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 19	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 20	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 21	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 22	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 23	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 24	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 25	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 26	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 27	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 28	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 29	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 30	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
1 31	24	.258-01	.151-02	.708	-1.11	-2.72	.285	.658-01	.300-01	.150	.120	.080-01	.04
MONTH	TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MID-RANGE	AV. AD. DEV
1 73	733	.243-01	.303-03	5.71	61.9	-5.53	10.1	.398-02	.100-06	.450	.450	.225	.02

ORIGINAL PAGE IS
 OF POOR QUALITY

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOX.

MONTH AND YEAR: 1 73

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999	-.3514-01	0
-.3514-01	-.2776-01	0
-.2776-01	-.2027-01	0
-.2027-01	-.1284-01	0
-.1284-01	-.5412-02	0
-.5412-02	.2019-02	100
.2019-02	.9450-02	0
.9450-02	.1080-01	222
.1080-01	.2431-01	178
.2431-01	.3174-01	94
.3174-01	.3917-01	0
.3917-01	.4660-01	49
.4660-01	.5403-01	26
.5403-01	.6146-01	20
.6146-01	.6890-01	0
.6890-01	.7633-01	13
.7633-01	.8376-01	7
.8376-01	.9999	24

CHI-SQUARED VALUE WITH 7 D. F. IS 300.61

TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS B= .01 AND C= 27.52

CHI-SQUARED VALUE WITH 7 D. F. IS 318.56

TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .02

CHI-SQUARED VALUE WITH 8 D. F. IS 157.21

FREQUENCY DISTRIBUTION FOR LOGS OF OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999	-14.00	100
-14.00	-12.90	0
-12.90	-11.92	0
-11.92	-10.86	0
-10.86	-9.78	0
-9.78	-8.72	0
-8.72	-7.66	0
-7.66	-6.59	0
-6.59	-5.53	0
-5.53	-4.46	222
-4.46	-3.40	272
-3.40	-2.33	16
-2.33	-1.27	1
-1.27	-.20	0
-.20	.85	0
.85	1.92	0
1.92	2.99	0
2.99	.9999	0

CHI-SQUARED VALUE WITH 20 D. F. IS 1200.50

MONTH AND YEAR: 1 73

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	APPLN BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-.4844-01	0
-.4844-01	-.4430-01	4
-.4230-01	-.3633-01	5
-.3633-01	-.3027-01	26
-.3027-01	-.2422-01	15
-.2422-01	-.1816-01	26
-.1816-01	-.1211-01	41
-.1211-01	-.0605-02	163
-.0605-02	-.1150-02	172
-.1150-02	-.0653-02	116
-.0653-02	-.1211-01	60
-.1211-01	-.1816-01	38
-.1816-01	-.2422-01	17
-.2422-01	-.3027-01	8
-.3027-01	-.3633-01	9
-.3633-01	-.4230-01	9
-.4230-01	-.4844-01	21
-.4844-01	9999.	

CHI-SQUARED VALUE WITH 120 D.F. IS 427.10

OBSERVATIONS ADJUSTED BY SUBTRACTING 1-HOURLY AVERAGES

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

[illegible]

CHL-S ADDED VALUE WITH 100% F. IS 334.75

AVERAGES FOR THE SAME TIME PERIOD DURING THE 10TH

[illegible]

ORIGINAL PAGE IS
OF POOR QUALITY

AIR POLLUTION ANALYSIS

STATE: 46 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOX.

MONTH AND YEAR: 2 73

DATE	TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MID-RANGE	MEDIAN
2/1	24	.375-02	.234-04	.516	-1.73	-11.8	31.1	.750-05	.100-06	.100-01	.100-01	.500-02	.00
2/2	23	.104-01	.737-04	1.57	3.79	-6.96	23.4	.952-03	.100-06	.400-01	.400-01	.200-01	.01
2/3	24	.146-01	.125-03	1.18	.738	-5.73	15.7	.326-02	.100-06	.400-01	.400-01	.200-01	.01
2/4	24	.708-02	.221-03	3.18	10.5	-12.1	32.2	.549-05	.100-06	.700-01	.700-01	.350-01	.00
2/5	24	.208-02	.246-04	2.38	4.79	-14.2	19.0	.701-06	.100-06	.200-01	.200-01	.100-01	.00
2/6	24	.208-02	.155-04	1.44	.632-01	-13.7	21.9	.110-05	.100-06	.100-01	.100-01	.500-02	.00
2/7	24	.125-01	.271-04	2.60	3.09	-4.44	.104	.118-01	.100-01	.300-01	.200-01	.200-01	.01
2/8	24	.154-01	.098-04	-1.03	-1.03	-6.53	24.3	.146-02	.100-06	.300-01	.300-01	.150-01	.02
2/9	24	.192-01	.150-04	-1.07	-2.71	-3.98	.620-01	.187-01	.100-01	.300-01	.200-01	.200-01	.02
2/10	24	.121-01	.419-04	-1.225	-1.686	-5.81	15.3	.299-02	.100-06	.200-01	.200-01	.100-01	.01
2/11	24	.150-01	.505-04	-1.56	-1.250	-5.61	15.3	.366-02	.100-06	.200-01	.200-01	.100-01	.02
2/12	24	.175-01	.771-04	-1.000	-1.250	-5.40	16.3	.415-02	.100-06	.300-01	.300-01	.150-01	.02
2/13	24	.275-01	.677-03	1.620	1.94	-3.94	.615	.194-01	.100-01	.110	.100-00	.600-01	.02
2/14	24	.108-01	.100-04	4.59	19.0	-4.56	.482-01	.105-01	.100-01	.300-01	.200-01	.200-01	.01
2/15	24	.275-01	.144-03	.035	-1.798	-3.69	.196	.250-01	.100-01	.500-01	.400-01	.300-01	.02
2/16	23	.226-01	.150-03	1.60	3.53	-3.93	.263	.197-01	.100-01	.500-01	.400-01	.300-01	.02
2/17	24	.135-01	.130-03	.123	-1.47	-8.13	33.6	.277-03	.100-06	.300-01	.300-01	.150-01	.01
2/18	24	.338-01	.463-03	.357	.211	-4.53	12.8	.108-01	.100-06	.900-01	.900-01	.450-01	.03
2/19	24	.208-01	.333-03	1.48	2.30	-5.48	18.5	.416-02	.100-06	.800-01	.800-01	.400-01	.02
2/20	24	.312-01	.270-03	-1.59	0.30	-3.75	.446	.235-01	.100-01	.150	.140	.800-01	.02
2/21	24	.425-01	.235-03	-1.779-01	-1.14	-3.23	.167	.394-01	.200-01	.700-01	.500-01	.450-01	.04
2/22	24	.290-01	.187-03	.503	1.95-01	-3.64	.271	.262-01	.100-01	.600-01	.500-01	.350-01	.03
2/23	24	.246-01	.101-03	1.50	2.69	-3.65	.284	.213-01	.100-01	.700-01	.600-01	.400-01	.02
2/24	24	.296-01	.440-03	1.49	1.57	-3.74	.424	.238-01	.100-01	.900-01	.800-01	.500-01	.02
2/25	24	.246-01	.135-03	.508	-1.244	-3.83	.265	.218-01	.100-01	.500-01	.400-01	.300-01	.02
2/26	24	.136-01	.130-03	.355	-1.818-01	-7.56	30.3	.522-03	.100-06	.400-01	.400-01	.200-01	.02
2/27	24	.100-01	.170-04	-1.25-12	.000	-16.1	-7.63-05	.100-06	.100-06	.100-06	.000	.100-06	.00
2/28	23	.505-02	.420-04	.717	-.525	-10.6	33.8	.262-04	.100-06	.200-01	.200-01	.100-01	.00
MONTH	TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MID-RANGE	AV. AB. DEV
2 73	668	.175-01	.294-03	2.08	7.99	-0.67	25.7	.127-02	.100-06	.150	.150	.750-01	.01

AIR POLLUTION ANALYSIS
 STATE: 48 VA.
 AREA: 0710 CHESAPEAKE
 SITE: 001 ARMORY
 POLLUTANT: 42401 SULFUR DIOX.
 MONTH AND YEAR: 7 73

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-1675-01	0
-1675-01	-1247-01	0
-1247-01	-6184-02	0
-6184-02	-3901-02	0
-3901-02	.3824-03	148
.3824-03	.4660-02	0
.4660-02	.8949-02	0
.8949-02	.1323-01	148
.1323-01	.1751-01	0
.1751-01	.2180-01	179
.2180-01	.2600-01	0
.2600-01	.3036-01	64
.3036-01	.3465-01	0
.3465-01	.3893-01	0
.3893-01	.4321-01	34
.4321-01	.4750-01	0
.4750-01	.5174-01	17
.5174-01	9999.	28

CHI-SQUARED VALUE WITH 40 D. F. IS 130.76
 TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS H= .05 AND C= 59.67
 CHI-SQUARED VALUE WITH 4 D. F. IS 1628.48
 TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .32
 CHI-SQUARED VALUE WITH 5 D. F. IS 1339.18

FREQUENCY DISTRIBUTION FOR LOGS OF OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-16.80	0
-16.80	-15.54	148
-15.54	-14.27	0
-14.27	-13.00	0
-13.00	-11.74	0
-11.74	-10.47	0
-10.47	-9.20	0
-9.20	-7.93	0
-7.93	-6.67	0
-6.67	-5.40	0
-5.40	-4.13	148
-4.13	-2.87	294
-2.87	-1.60	28
-1.60	-.33	0
-.33	.94	0
.94	2.19	0
2.19	3.46	0
3.46	9999.	0

CHI-SQUARED VALUE WITH 15 D. F. IS 1751.00

ORIGINAL PAGE IS
 OF POOR QUALITY

AIR POLLUTION ANALYSIS
 STATE: 46 VA.
 AREA: 0710 CHESAPEAKE
 SITE: 001 ARMORY
 POLLUTANT: 42401 SULFUR DIOX.
 MONTH AND YEAR: 2 73

OBSERVATIONS ADJUSTED BY SUBTRACTING DAILY AVERAGES

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-2694-01	2
-2694-01	-2350-01	4
-2350-01	-2021-01	12
-2021-01	-1684-01	27
-1684-01	-1347-01	39
-1347-01	-1010-01	51
-1010-01	-673-02	51
-673-02	-336-02	60
-336-02	749-09	158
749-09	336-02	56
336-02	673-02	86
673-02	1010-01	29
1010-01	1347-01	17
1347-01	1684-01	14
1684-01	2021-01	13
2021-01	2350-01	6
2350-01	2694-01	10
2694-01	9999.	25

CHI-SQUARED VALUE WITH 130 D.F. IS 225.59

OBSERVATIONS ADJUSTED BY SUBTRACTING 1-HOURLY AVERAGES

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-334-01	0
-334-01	-2927-01	0
-2927-01	-2504-01	0
-2504-01	-2091-01	30
-2091-01	-1673-01	20
-1673-01	-1254-01	109
-1254-01	-836-02	53
-836-02	-4181-02	70
-4181-02	-255-02	129
-255-02	4181-02	55
4181-02	836-02	73
836-02	1254-01	18
1254-01	1673-01	21
1673-01	2091-01	31
2091-01	2504-01	7
2504-01	2927-01	14
2927-01	334-01	3
334-01	9999.	55

CHI-SQUARED VALUE WITH 110 D.F. IS 242.14

AVERAGES FOR THE SAME TIME EACH DAY DURING THE MONTH

OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER
0150	0124	0161	0144	0134	0144	0161	0164	0232	0239	0193	0207				
0222	0221	0232	0196	0111	0204	0161	0164	0121	0139	0132	0136				

ORIGINAL PAGE IS
 OF POOR QUALITY

AIR POLLUTION ANALYSIS

STATE: 46 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42901 SULFUR DIOX.

MONTH AND YEAR: 3 73

DATE	TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MID-RANGE	AV. AB. DEV
3 1	24	.204-01	.721-03	1.05	1.80	-8.38	36.3	.230-03	.100-06	.100+00	.100+00	.500-01	.01
3 2	24	.268-01	.179-02	2.69	7.35	-7.17	30.7	.769-03	.100-06	.190	.190	.950-01	.01
3 3	24	.633-03	.764-05	3.02	7.69	-15.2	10.1	.261-06	.100-06	.100-01	.100-01	.500-02	.00
3 4	24	.100-01	.342-03	2.18	3.50	-11.1	35.4	.150-04	.100-06	.700-01	.700-01	.350-01	.00
3 5	24	.417-03	.399-05	4.59	19.0	-15.6	5.29	.162-06	.100-06	.100-01	.100-01	.500-02	.00
3 6	24	.125-02	.359-04	4.59	19.0	-15.6	33.5	.169-06	.100-06	.300-01	.300-01	.150-01	.00
3 7	24	.542-02	.332-04	.492	-7.11	-10.3	3.25	.144-01	.100-06	.500-01	.500-01	.250-01	.03
3 8	24	.207-01	.189-03	-1.57	-1.986	-4.24	5.45	.144-01	.100-06	.200-01	.200-01	.100-01	.00
3 9	24	.503-02	.326-04	.341	-7.85	-9.85	33.2	.526-04	.100-06	.200-01	.200-01	.100-01	.00
3 10	24	.102-02	.240-04	2.71	6.37	-14.5	16.3	.496-06	.100-06	.200-01	.200-01	.100-01	.00
3 11	24	.833-02	.164-03	1.98	3.39	-10.7	35.4	.235-04	.100-06	.500-01	.500-01	.250-01	.00
3 12	24	.196-01	.214-03	1.96	3.06	-4.13	34.0	.161-01	.100-01	.700-01	.600-01	.400-01	.01
3 13	24	.100-01	.174-03	2.82	5.27	-9.13	35.1	.100-03	.100-06	.600-01	.600-01	.300-01	.01
3 14	24	.539-01	.772-03	.653	-1.614	-4.57	13.0	.104-01	.100-06	.800-01	.800-01	.400-01	.03
3 15	24	.204-01	.154-03	1.36	1.89	-4.05	31.2	.439-01	.100-01	.120	.110	.650-01	.03
3 16	24	.121-01	.415-04	.711	1.61	-5.37	10.6	.463-02	.100-01	.600-01	.500-01	.350-01	.02
3 17	24	.333-02	.114-03	3.63	12.5	-14.6	16.5	.464-06	.100-06	.300-01	.300-01	.150-01	.01
3 18	24	.154-01	.582-04	1.55	2.25	-4.27	17.9	.140-01	.100-01	.500-01	.500-01	.250-01	.00
3 19	24	.209-01	.142-02	2.93	8.26	-8.20	34.0	.276-03	.100-06	.170	.170	.850-01	.01
3 20	24	.625-02	.651-04	.761	-1.02	-11.2	34.2	.140-04	.100-06	.200-01	.200-01	.100-01	.00
3 21	24	.148-01	.130-03	-1.09	-1.50	-6.10	34.3	.304-03	.100-06	.300-01	.300-01	.150-01	.02
3 22	24	.200-02	.142-03	.146	-1.05	-9.42	16.6	.444-02	.100-06	.400-01	.400-01	.200-01	.02
3 23	24	.179-01	.450-03	.772	-1.924	-9.35	30.1	.873-04	.100-06	.600-01	.600-01	.300-01	.01
3 24	24	.146-01	.704-03	1.42	-1.262	-12.4	32.4	.273-05	.100-06	.800-01	.800-01	.400-01	.00
3 25	24	.500-02	.243-04	.338	-1.248	-11.3	33.0	.124-04	.100-06	.200-01	.200-01	.100-01	.00
3 26	24	.417-02	.109-04	2.27	3.14	-14.7	14.5	.121-04	.100-06	.100-01	.100-01	.500-02	.00
3 27	24	.125-02	.165-04	1.44	.632-01	-13.7	21.4	.422-06	.100-06	.100-01	.100-01	.500-02	.00
3 28	24	.208-02						.110-05	.100-06	.100-01	.100-01	.500-02	.00
3 29	NO OBSERVATIONS RECORDED.												
3 30	NO OBSERVATIONS RECORDED.												
3 31	NO OBSERVATIONS RECORDED.												
3 73	689	.131-01	.417-03	3.13	15.7	-9.59	36.8	.686-04	.100-06	.190	.190	.950-01	.01

ORIGINAL PAGE IS
OF POOR QUALITY

AIR POLLUTION ANALYSIS

STATE: 46 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMOY

POLLUTANT: 42401 SULFUR DIOX.

MONTH AND YEAR: 3 74

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
9999.	2771-01	0
2771-01	2661-01	0
2661-01	1750-01	0
1750-01	1640-01	0
1640-01	7290-02	0
7290-02	2192-02	0
2192-02	2912-02	318
2912-02	6010-02	0
6010-02	1512-01	176
1512-01	1822-01	0
1822-01	2333-01	74
2333-01	2043-01	0
2043-01	3354-01	54
3354-01	3864-01	0
3864-01	4374-01	13
4374-01	4685-01	0
4685-01	5395-01	14
5395-01	9999.	35

CHI-SQUARED VALUE WITH 40 D. F. IS 180.84

TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS $\mu = .01$ AND $C = 31.48$

CHI-SQUARED VALUE WITH 40 D. F. IS 1371.24

TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .01

CHI-SQUARED VALUE WITH 5 D. F. IS 372.95

FREQUENCY DISTRIBUTION FOR LOGS OF OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
9999.	21.72	0
21.72	20.21	0
20.21	18.69	0
18.69	17.17	0
17.17	15.65	318
15.65	14.14	0
14.14	12.62	0
12.62	11.10	0
11.10	9.587	0
9.587	8.079	0
8.079	6.553	0
6.553	5.030	0
5.030	3.519	250
3.519	2.002	119
2.002	1.4847	2
1.4847	1.032	0
1.032	2.549	0
2.549	9999.	0

NOT ENOUGH OBSERVATIONS TO CALCULATE CHI SQUARED STATISTIC.

AIR POLLUTION ANALYSIS
 STATE: 48 VA.
 AREA: 0710 CHESAPEAKE
 SITE: 001 ARMORY
 POLLUTANT: 42401 SULFUR DIOX.
 MONTH AND YEAR: 3 73

OBSERVATIONS ADJUSTED BY SUBTRACTING DAILY AVERAGES

FREQUENCY DISTRIBUTION FOR OBSERVATIONS		
LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-.9999	-.3370-01	3
-.3370-01	-.2949-01	3
-.2949-01	-.2527-01	7
-.2527-01	-.2106-01	8
-.2106-01	-.1685-01	31
-.1685-01	-.1264-01	46
-.1264-01	-.0842-02	71
-.0842-02	-.0421-02	57
-.0421-02	.5663-09	212
.5663-09	.4212-02	38
.4212-02	.0842-02	61
.0842-02	.1264-01	40
.1264-01	.1685-01	25
.1685-01	.2106-01	16
.2106-01	.2527-01	4
.2527-01	.2949-01	4
.2949-01	.3370-01	28
.3370-01	.9999	

CHI-SQUARED VALUE WITH 110 D.F. IS 414.41

OBSERVATIONS ADJUSTED BY SUBTRACTING 1-HOURLY AVERAGES

FREQUENCY DISTRIBUTION FOR OBSERVATIONS		
LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-.9999	-.4905-01	0
-.4905-01	-.3504-01	0
-.3504-01	-.3004-01	0
-.3004-01	-.2503-01	0
-.2503-01	-.2003-01	21
-.2003-01	-.1502-01	25
-.1502-01	-.1001-01	135
-.1001-01	-.5000-02	137
-.5000-02	.2061-08	99
.2061-08	.5060-02	60
.5060-02	.1001-01	36
.1001-01	.1502-01	29
.1502-01	.2003-01	24
.2003-01	.2503-01	19
.2503-01	.3004-01	1
.3004-01	.3504-01	9
.3504-01	.4005-01	36
.4005-01	.9999	

CHI-SQUARED VALUE WITH 100 D.F. IS 436.82

AVERAGES FOR THE SAME TIME EACH DAY DURING THE MONTH

00SERV NUMBER 1	00SERV NUMBER 2	00SERV NUMBER 3	00SERV NUMBER 4	00SERV NUMBER 5	00SERV NUMBER 6	00SERV NUMBER 7	00SERV NUMBER 8	00SERV NUMBER 9	00SERV NUMBER 10	00SERV NUMBER 11	00SERV NUMBER 12
.0103	.0064	.0083	.0131	.0134	.0145	.0128	.0141	.0186	.0221	.0169	.0155
00SERV NUMBER 13	00SERV NUMBER 14	00SERV NUMBER 15	00SERV NUMBER 16	00SERV NUMBER 17	00SERV NUMBER 18	00SERV NUMBER 19	00SERV NUMBER 20	00SERV NUMBER 21	00SERV NUMBER 22	00SERV NUMBER 23	00SERV NUMBER 24
.0157	.0204	.0174	.0138	.0141	.0114	.0093	.0097	.0107	.0107	.0072	.0086

ORIGINAL PAGE IS
OF POOR QUALITY

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOX.

MONTH AND YEAR: 4 73

DATE	TALLY	MEAN	VARIANCE	COEFF. SKENNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	STD-RANGE	MEDIAN
4 1	NO OBSERVATIONS RECORDED.												
4 2	NO OBSERVATIONS RECORDED.												
4 3	NO OBSERVATIONS RECORDED.												
4 4	NO OBSERVATIONS RECORDED.												
4 5	NO OBSERVATIONS RECORDED.												
4 6	NO OBSERVATIONS RECORDED.												
4 7	NO OBSERVATIONS RECORDED.												
4 8	NO OBSERVATIONS RECORDED.												
4 9	NO OBSERVATIONS RECORDED.												
4 10	NO OBSERVATIONS RECORDED.												
4 11	NO OBSERVATIONS RECORDED.												
4 12	NO OBSERVATIONS RECORDED.												
4 13	NO OBSERVATIONS RECORDED.												
4 14	NO OBSERVATIONS RECORDED.												
4 15	NO OBSERVATIONS RECORDED.												
4 16	NO OBSERVATIONS RECORDED.												
4 17	NO OBSERVATIONS RECORDED.												
4 18	NO OBSERVATIONS RECORDED.												
4 19	NO OBSERVATIONS RECORDED.												
4 20	NO OBSERVATIONS RECORDED.												
4 21	NO OBSERVATIONS RECORDED.												
4 22	NO OBSERVATIONS RECORDED.												
4 23	NO OBSERVATIONS RECORDED.												
4 24	NO OBSERVATIONS RECORDED.												
4 25	NO OBSERVATIONS RECORDED.												
4 26	NO OBSERVATIONS RECORDED.												
4 27	NO OBSERVATIONS RECORDED.												
4 28	NO OBSERVATIONS RECORDED.												
4 29	NO OBSERVATIONS RECORDED.												
4 30	NO OBSERVATIONS RECORDED.												

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOX.

ANNUAL SUMMARY: 1973

TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	TD-RANGE
2090	.184-01	.563-03	5.07	61.0	-7.23	20.6	.723-03	.100-06	.450	.450	.225

ORIGINAL PAGE IS
OF POOR QUALITY

ANNUAL SUMMARY: 1973

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-.9999	-.2900-.01	0
-.2900-.01	-.2500-.01	0
-.2500-.01	-.1715-.01	0
-.1715-.01	-.1120-.01	0
-.1120-.01	-.0573-.03	0
-.0573-.03	.0573-.03	596
.0573-.03	.0573-.03	0
.0573-.03	.1252-.01	596
.1252-.01	.1643-.01	0
.1643-.01	.2433-.01	501
.2433-.01	.3631-.01	212
.3631-.01	.4217-.01	0
.4217-.01	.4217-.01	101
.4217-.01	.4410-.01	0
.4410-.01	.5410-.01	0
.5410-.01	.5990-.01	0
.5990-.01	.6590-.01	45
.6590-.01	.9999	42

CHI-SQUARED VALUE WITH 6 D. F. IS 3558.35

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-10.10	0
-10.10	-10.70	0
-10.70	-15.40	596
-15.40	-14.00	0
-14.00	-12.80	0
-12.80	-11.30	0
-11.30	-9.95	0
-9.95	-8.59	0
-8.59	-7.63	0
-7.63	-5.67	0
-5.67	-4.51	596
-4.51	-3.14	744
-3.14	-1.78	178
-1.78	-1.42	6
-1.42	.93	0
.93	2.29	0
2.29	3.65	0
3.65	9999.	0

OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER	OBSERV NUMBER
1	2	3	4	5	6	7	8	9	10	11	12	13

.0162 .0134 .0154 .0165 .0165 .0169 .0157 .0154 .0144 .0200 .0189 .0159
OPSRV OPSRV OPSRV OPSRV OPSRV OPSRV OPSRV OPSRV OPSRV OPSRV OPSRV OPSRV OPSRV
NUMBER 13 NUMBER 14 NUMBER 15 NUMBER 16 NUMBER 17 NUMBER 18 NUMBER 19 NUMBER 20 NUMBER 21 NUMBER 22 NUMBER 23 NUMBER 24

.0179 .0187 .0183 .0183 .0175 .0185 .0203 .0202 .0153 .0160 .0156 .0157

ORIGINAL PAGE IS
OF POOR QUALITY

FREQUENCY DISTRIBUTION FOR DAILY MEANS FOR THE PERIOD 11 1 72 TO 4 30 73

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-1029-01	0
-1029-01	-9937-02	0
-9937-02	-3412-02	0
-3412-02	-1281-04	2
-1281-04	-3430-02	13
-3430-02	-9960-02	18
-9960-02	-1029-01	18
-1029-01	-1571-01	16
-1571-01	-1714-01	25
-1714-01	-2050-01	16
-2050-01	-2399-01	6
-2399-01	-2741-01	9
-2741-01	-3084-01	7
-3084-01	-3428-01	6
-3428-01	-3769-01	1
-3769-01	-4111-01	2
-4111-01	-4454-01	1
-4454-01	9999.	9

CHI-SQUARED VALUE WITH 8D. F. IS 20.89

TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS B= .56 AND C= 91.31

CHI-SQUARED VALUE WITH 7 D. F. IS 9.96

TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .02

CHI-SQUARED VALUE WITH 8 D. F. IS 26.74

FREQUENCY DISTRIBUTION FOR LOGS OF OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-7.902	2
-7.902	-7.486	2
-7.486	-7.070	3
-7.070	-6.653	3
-6.653	-6.237	4
-6.237	-5.821	8
-5.821	-5.405	11
-5.405	-4.989	18
-4.989	-4.572	30
-4.572	-4.156	33
-4.156	-3.740	23
-3.740	-3.324	9
-3.324	-2.907	3
-2.907	-2.491	0
-2.491	-2.075	0
-2.075	-1.659	0
-1.659	-1.243	0
-1.243	9999.	0

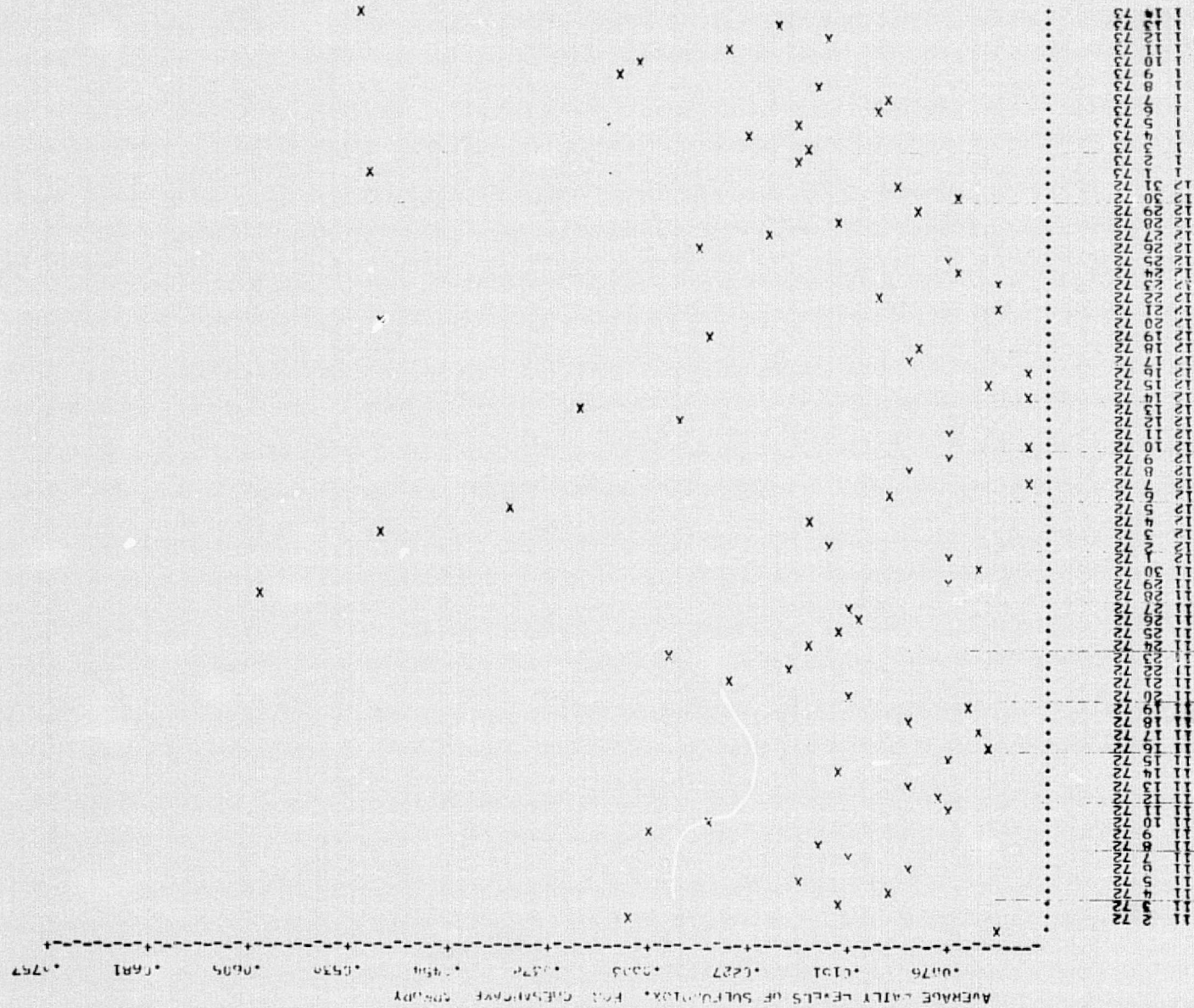
CHI-SQUARED VALUE WITH 4D. F. IS 65.98

Example 2

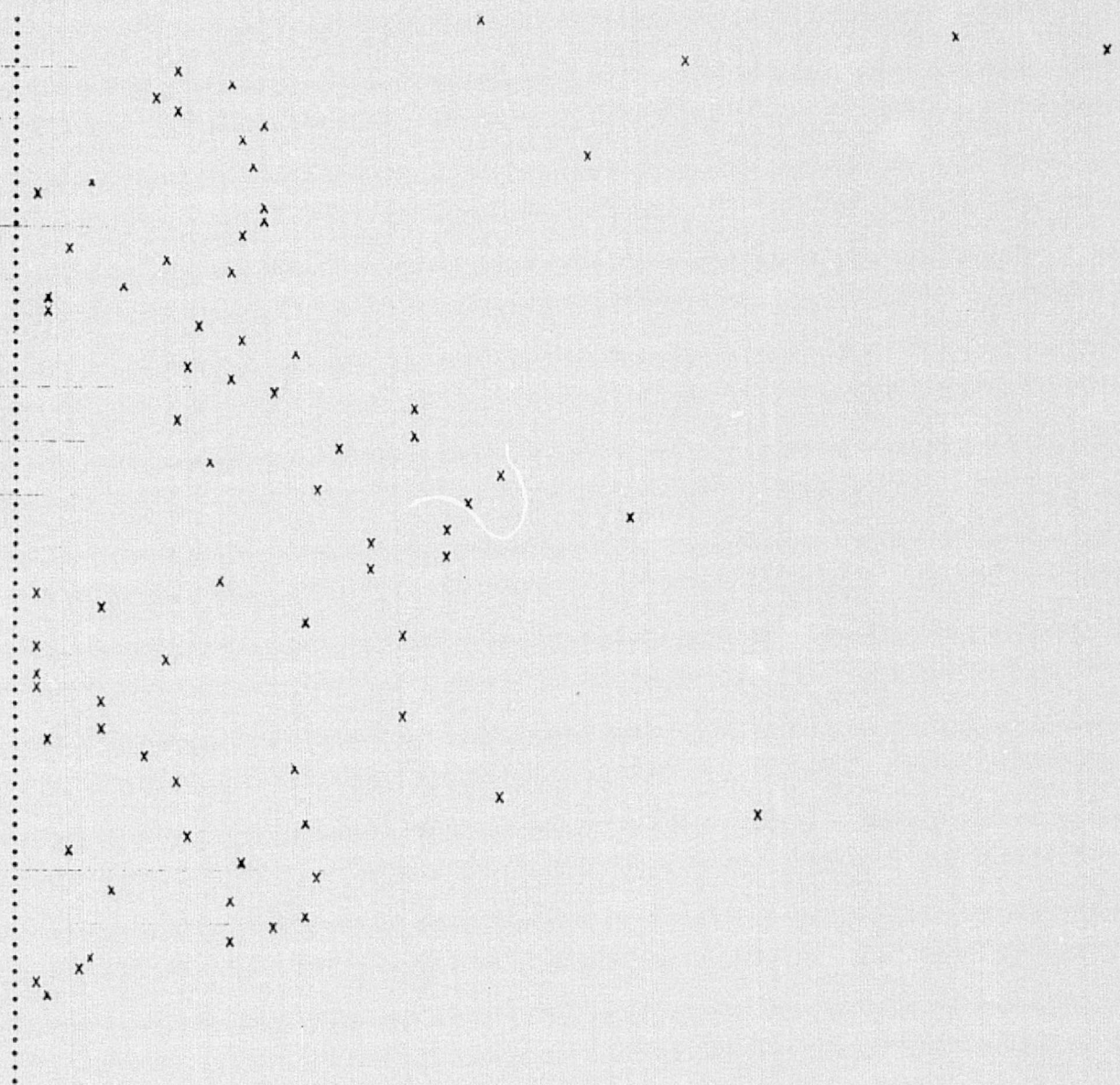
PERIODIC AVERAGES FOR PERIODS OF 7 DAYS FOR DATA FROM 11 1 76 TO 4 30 75

AVERAGE FOR WED.	.0229
AVERAGE FOR THR.	.0203
AVERAGE FOR FRI.	.0202
AVERAGE FOR SAT.	.0151
AVERAGE FOR SUN.	.0111
AVERAGE FOR MON.	.0151
AVERAGE FOR TUE.	.0149

ORIGINAL PAGE IS
OF POOR QUALITY



1 15 73
1 16 73
1 17 73
1 18 73
1 19 73
1 20 73
1 21 73
1 22 73
1 23 73
1 24 73
1 25 73
1 26 73
1 27 73
1 28 73
1 29 73
1 30 73
1 31 73
2 1 73
2 2 73
2 3 73
2 4 73
2 5 73
2 6 73
2 7 73
2 8 73
2 9 73
2 10 73
2 11 73
2 12 73
2 13 73
2 14 73
2 15 73
2 16 73
2 17 73
2 18 73
2 19 73
2 20 73
2 21 73
2 22 73
2 23 73
2 24 73
2 25 73
2 26 73
2 27 73
2 28 73
2 29 73
2 30 73
2 31 73
3 1 73
3 2 73
3 3 73
3 4 73
3 5 73
3 6 73
3 7 73
3 8 73
3 9 73
3 10 73
3 11 73
3 12 73
3 13 73
3 14 73
3 15 73
3 16 73
3 17 73
3 18 73
3 19 73
3 20 73
3 21 73
3 22 73
3 23 73
3 24 73
3 25 73
3 26 73
3 27 73
3 28 73
3 29 73
3 30 73
3 31 73
4 1 73
4 2 73
4 3 73
4 4 73
4 5 73
4 6 73
4 7 73
4 8 73
4 9 73
4 10 73
4 11 73
4 12 73
4 13 73
4 14 73
4 15 73
4 16 73
4 17 73
4 18 73
4 19 73
4 20 73
4 21 73
4 22 73
4 23 73
4 24 73
4 25 73
4 26 73
4 27 73
4 28 73
4 29 73
4 30 73
4 31 73
5 1 73
5 2 73
5 3 73
5 4 73
5 5 73
5 6 73
5 7 73
5 8 73
5 9 73
5 10 73
5 11 73
5 12 73
5 13 73
5 14 73
5 15 73
5 16 73
5 17 73
5 18 73
5 19 73
5 20 73
5 21 73
5 22 73
5 23 73
5 24 73
5 25 73
5 26 73
5 27 73
5 28 73
5 29 73
5 30 73
5 31 73
6 1 73
6 2 73
6 3 73
6 4 73
6 5 73
6 6 73
6 7 73
6 8 73
6 9 73
6 10 73
6 11 73
6 12 73
6 13 73
6 14 73
6 15 73
6 16 73
6 17 73
6 18 73
6 19 73
6 20 73
6 21 73
6 22 73
6 23 73
6 24 73
6 25 73
6 26 73
6 27 73
6 28 73
6 29 73
6 30 73
6 31 73
7 1 73
7 2 73
7 3 73
7 4 73
7 5 73
7 6 73
7 7 73
7 8 73
7 9 73
7 10 73
7 11 73
7 12 73
7 13 73
7 14 73
7 15 73
7 16 73
7 17 73
7 18 73
7 19 73
7 20 73
7 21 73
7 22 73
7 23 73
7 24 73
7 25 73
7 26 73
7 27 73
7 28 73
7 29 73
7 30 73
7 31 73
8 1 73
8 2 73
8 3 73
8 4 73
8 5 73
8 6 73
8 7 73
8 8 73
8 9 73
8 10 73
8 11 73
8 12 73
8 13 73
8 14 73
8 15 73
8 16 73
8 17 73
8 18 73
8 19 73
8 20 73
8 21 73
8 22 73
8 23 73
8 24 73
8 25 73
8 26 73
8 27 73
8 28 73
8 29 73
8 30 73
8 31 73
9 1 73
9 2 73
9 3 73
9 4 73
9 5 73
9 6 73
9 7 73
9 8 73
9 9 73
9 10 73
9 11 73
9 12 73
9 13 73
9 14 73
9 15 73
9 16 73
9 17 73
9 18 73
9 19 73
9 20 73
9 21 73
9 22 73
9 23 73
9 24 73
9 25 73
9 26 73
9 27 73
9 28 73
9 29 73
9 30 73
9 31 73



ORIGINAL PAGE IS
OF POOR QUALITY

6	3	73	.
7	0	73	.
8	7	73	.
9	8	73	.
10	9	73	.
11	10	73	.
12	11	73	.
13	12	73	.
14	13	73	.
15	14	73	.
16	15	73	.
17	16	73	.
18	17	73	.
19	18	73	.
20	19	73	.
21	20	73	.
22	21	73	.
23	22	73	.
24	23	73	.
25	24	73	.
26	25	73	.
27	26	73	.
28	27	73	.
29	28	73	.
30	29	73	.
31	30	73	.

Example 3

FOLLOWING ARE DATA RECORD IMAGES (WITH SOME SPACING PROVIDED).

[illegible]

ORIGINAL PAGE IS
OF POOR QUALITY

480710001F021721210124240114072	0.	0.	0.	0.	0.	0.	1.	1.	0.	0.	0.
480710001F02172121104240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721211124240114072	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172121204240114072	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721212124240114072	0.	9999.	9999.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172121304240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721213124240114072	2.	1.	2.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172121404240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721214124240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172121504240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721215124240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172121604240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721216124240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172121704240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721217124240114072	2.	1.	3.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172121804240114072	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721218124240114072	1.	9999.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172121904240114072	1.	1.	2.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721219124240114072	2.	2.	2.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172122004240114072	2.	3.	11.	12.	1.	17.	13.	11.	2.	2.	2.
480710001F021721220124240114072	4.	4.	14.	3.	0.	2.	2.	2.	2.	1.	1.
480710001F02172122104240114072	1.	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721221124240114072	1.	1.	1.	1.	0.	0.	0.	0.	0.	0.	0.
480710001F02172122204240114072	0.	2.	2.	2.	3.	1.	1.	0.	0.	0.	0.
480710001F021721222124240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172122304240114072	0.	0.	1.	1.	0.	0.	0.	0.	0.	0.	0.
480710001F021721223124240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172122404240114072	2.	1.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721224124240114072	0.	4.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172122504240114072	1.	1.	1.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721225124240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F02172122604240114072	1.	2.	1.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721226124240114072	4.	4.	3.	3.	0.	4.	4.	9999.	9999.	9999.	9999.
480710001F02172122704240114072	9999.	9999.	9999.	9999.	9999.	9999.	9999.	9999.	9999.	9999.	9999.
480710001F021721227124240114072	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.
480710001F02172122804240114072	3.	1.	2.	2.	1.	1.	1.	1.	1.	1.	1.
480710001F021721228124240114072	1.	2.	1.	1.	1.	1.	1.	1.	1.	1.	1.
480710001F02172122904240114072	1.	1.	2.	1.	1.	1.	1.	1.	1.	1.	1.
480710001F021721229124240114072	0.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.
480710001F02172123004240114072	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
480710001F021721230124240114072	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.
480710001F02172123104240114072	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.
480710001F021721231124240114072	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	3.

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOX.

ANNUAL SUMMARY: 1972

MONTH	TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	STD-RANGE	AV. AB. DEV
11 72	710	.159-01	.315-03	2.55	9.32	-0.86	25.8	.105-02	.100-06	.140	.140	.700-01	.01
12 72	727	.145-01	.903-03	7.16	89.2	-9.04	38.5	.110-03	.100-06	.460	.460	.230	.01

ORIGINAL PAGE IS
OF POOR QUALITY

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 0711 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOX.

ANNUAL SUMMARY: 1972

TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	MID-RANGE
1437	.152-01	.562-03	6.07	93.4	-7.96	31.9	.348-03	.100-06	.460	.460	.230

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 02401 SULFUR DIOX.

ANNUAL SUMMARY: 1972

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-3221-01	0
-3221-01	-2029-01	0
-2029-01	-2030-01	0
-2030-01	-1443-01	0
-1443-01	-8501-02	0
-8501-02	-2573-02	0
-2573-02	3350-02	403
3350-02	9284-02	531
9284-02	1521-01	211
1521-01	2114-01	0
2114-01	2707-01	0
2707-01	3300-01	73
3300-01	3693-01	0
3693-01	4485-01	05
4485-01	5078-01	29
5078-01	5671-01	0
5671-01	6264-01	14
6264-01	9999.	51

CHI-SQUARED VALUE WITH 50 D.F. IS 524.18

TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS B= .01 AND C= 27.05

CHI-SQUARED VALUE WITH 5 D.F. IS 1576.47

TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .02

CHI-SQUARED VALUE WITH 6 D.F. IS 257.37

FREQUENCY DISTRIBUTION FOR LOGS OF OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-19.20	0
-19.20	-17.83	0
-17.83	-16.44	0
-16.44	-15.03	403
-15.03	-13.61	0
-13.61	-12.20	0
-12.20	-10.79	0
-10.79	-9.377	0
-9.377	-7.964	0
-7.964	-6.552	0
-6.552	-5.140	0
-5.140	-3.726	742
-3.726	-2.310	208
-2.310	-.9033	23
-.9033	0.5089	1
0.5089	1.921	0
1.921	3.333	0
3.333	9999.	0

CHI-SQUARED VALUE WITH 10 D.F. IS 849.99

ORIGINAL PAGE IS
OF POOR QUALITY

PROLOGUE.

1.
2.
3.
1.
1.
6.
1.
2.
1.
1.
1.
2.
1.
3.
7.
2.
1.
3.
2.
1.
2.
1.
2.
3.
9.
4.
8.
9.
4.
2.
5.
1.
2.
2.
0.
0.
0.
0.
1.
1.
2.
2.
0.
2.
5.
7.
1.
0.
1.
0.
0.
1.
1.
1.
1.
0.
0.
0.
0.
0.
1.
1.
1.
1.
2.
2.

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

480710001F02173 430124240114072 3999. 0999. 0999. 0999. 0999. 0999. 0999. 0999. 0999. 0999.

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 42401 SULFUR DIOX.

ANNUAL SUMMARY: 1973

MONTH	TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	STD-RANGE	AV. AB. DEV
1	73	733	.243-01	.883-03	5.71	-5.53	18.1	.396-02	.100-06	.450	.450	.225	.02
2	73	666	.175-01	.294-03	2.68	-6.67	25.7	.127-02	.100-06	.150	.150	.750-01	.01
3	73	689	.131-01	.417-03	3.13	-9.50	36.8	.646-04	.100-06	.190	.190	.950-01	.01

AIR POLLUTION ANALYSIS

STATE: 48 VA.

AREA: 0710 CHESAPEAKE

SITE: 001 ARMORY

POLLUTANT: 02901 SULFUR DIOX.

ANNUAL SUMMARY: 1973

TALLY	MEAN	VARIANCE	COEFF. SKEWNESS	COEFF. KURTOSIS	LOG MEAN	LOG VARIANCE	GEOMET. MEAN	MINIMUM	MAXIMUM	RANGE	STD-RANGE
2090	.184-01	.503-03	5.07	61.0	-7.23	20.6	.723-03	.100-06	.450	.450	.225

ORIGINAL PAGE IS
OF POOR QUALITY

AIR POLLUTION ANALYSIS
 STATE: 46 VA.
 AREA: 0710 CHESAPEAKE
 SITE: 001 ARMORY
 POLLUTANT: 42401 SULFUR DIOX.
 ANNUAL SUMMARY: 1973

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-2900-01	0
-2900-01	-2900-01	0
-2900-01	-1713-01	0
-1713-01	-1129-01	0
-1129-01	-5673-C2	0
-5673-C2	.6578-03	596
.6578-03	.6588-02	0
.6588-02	.1252-01	596
.1252-01	.1043-01	0
.1043-01	.2433-01	451
.2433-01	.3031-01	212
.3031-01	.3024-01	0
.3024-01	.4217-01	101
.4217-01	.4610-01	0
.4610-01	.5403-01	57
.5403-01	.5990-01	0
.5990-01	.6590-01	45
.6590-01	9999.	62

CHI-SQUARED VALUE WITH 50 D.F. IS 453.91

TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS H= .01 AND C= 32.78

CHI-SQUARED VALUE WITH 5 D.F. IS 7001.53

TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .02

CHI-SQUARED VALUE WITH 6 D.F. IS 3558.36

FREQUENCY DISTRIBUTION FOR LOGS OF OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-18.12	0
-18.12	-16.70	0
-16.70	-15.40	596
-15.40	-14.04	0
-14.04	-12.60	0
-12.60	-11.32	0
-11.32	-9.955	0
-9.955	-8.594	0
-8.594	-7.233	0
-7.233	-5.871	0
-5.871	-4.510	596
-4.510	-3.149	744
-3.149	-1.780	178
-1.780	-.4269	6
-.4269	.9384	0
.9384	2.295	0
2.295	3.650	0
3.650	9999.	0

CHI-SQUARED VALUE WITH 20 D.F. IS 3713.39

AVERAGE FOR WED.	.0029
AVERAGE FOR THU.	.0003
AVERAGE FOR FRI.	.0002
AVERAGE FOR SAT.	.0151
AVERAGE FOR SUN.	.0111
AVERAGE FOR MON.	.0051
AVERAGE FOR TUE.	.0149

AVERAGE FOR THR. .0403

AVERAGE FOR FBI. .01402

AVERAGE FOR SAT.	.0151
------------------	-------

AVERAGE FOR SUN. .0111

AVERAGE FOR MUN. .0451

AVERAGE FOR TUE. .0149

AVERAGES FOR THE SAME TIME EACH DAY DURING THE PERIOD 11 1 72 TO 4 30 73

AVERAGES FOR THE SAME TIME EACH DAY DURING THE PERIOD 11 1 72 TO 4 30 73
 OBSERV OBSERV OBSERV OBSERV OBSERV OBSERV OBSERV OBSERV OBSERV OBSERV
 NUMBER 1 NUMBER 2 NUMBER 3 NUMBER 4 NUMBER 5 NUMBER 6 NUMBER 7 NUMBER 8 NUMBER 9 NUMBER 10 NUMBER 11 NUMBER 12
 .0162 .0139 .0154 .0105 .0105 .0166 .0157 .0155 .0184 .0200 .0180 .0159
 OBSERV OBSERV OBSERV OBSERV OBSERV OBSERV OBSERV OBSERV OBSERV OBSERV
 NUMBER 13 NUMBER 14 NUMBER 15 NUMBER 16 NUMBER 17 NUMBER 18 NUMBER 19 NUMBER 20 NUMBER 21 NUMBER 22 NUMBER 23 NUMBER 24
 .0179 .0187 .0183 .0103 .0175 .0185 .0203 .0202 .0153 .0160 .0156 .0157

ORIGINAL PAGE IS
OF POOR QUALITY

FREQUENCY DISTRIBUTION FOR DAILY MEANS FOR THE PERIOD 11 1 72 10 4 30 73

FREQUENCY DISTRIBUTION FOR OBSERVATIONS

LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	..1020-01	0
..1020-n1	..6837-02	0
..6837-n2	..3412-02	0
..3412-n2	..1481-04	2
..1281-n4	..3433-02	13
..3434-n2	..0063-02	14
..6863-n2	..1029-01	18
..1029-n1	..1971-01	16
..1371-n1	..1714-01	25
..1714-n1	..2050-01	16
..2050-n1	..3399-01	6
..2399-n1	..2741-01	6
..2741-n1	..3084-01	7
..3084-n1	..3420-01	6
..3420-n1	..3769-01	1
..3769-n1	..4111-01	2
..4111-n1	..4454-01	1
..4454-n1	9999.	9

CHI-SQUARED VALUE WITH 80 D. F. IS 20.89

TEST FOR FIT TO GAMMA DISTRIBUTION WITH PARAMETERS $\alpha = .56$ AND $\beta = 91.31$

CHI-SQUARED VALUE WITH 7 D. F. IS 9.96

TEST FOR FIT TO EXPONENTIAL DISTRIBUTION WITH MEAN .02

CHI-SQUARED VALUE WITH 8 D. F. IS 20.74

FREQUENCY DISTRIBUTION FOR LOGS OF OBSERVATIONS

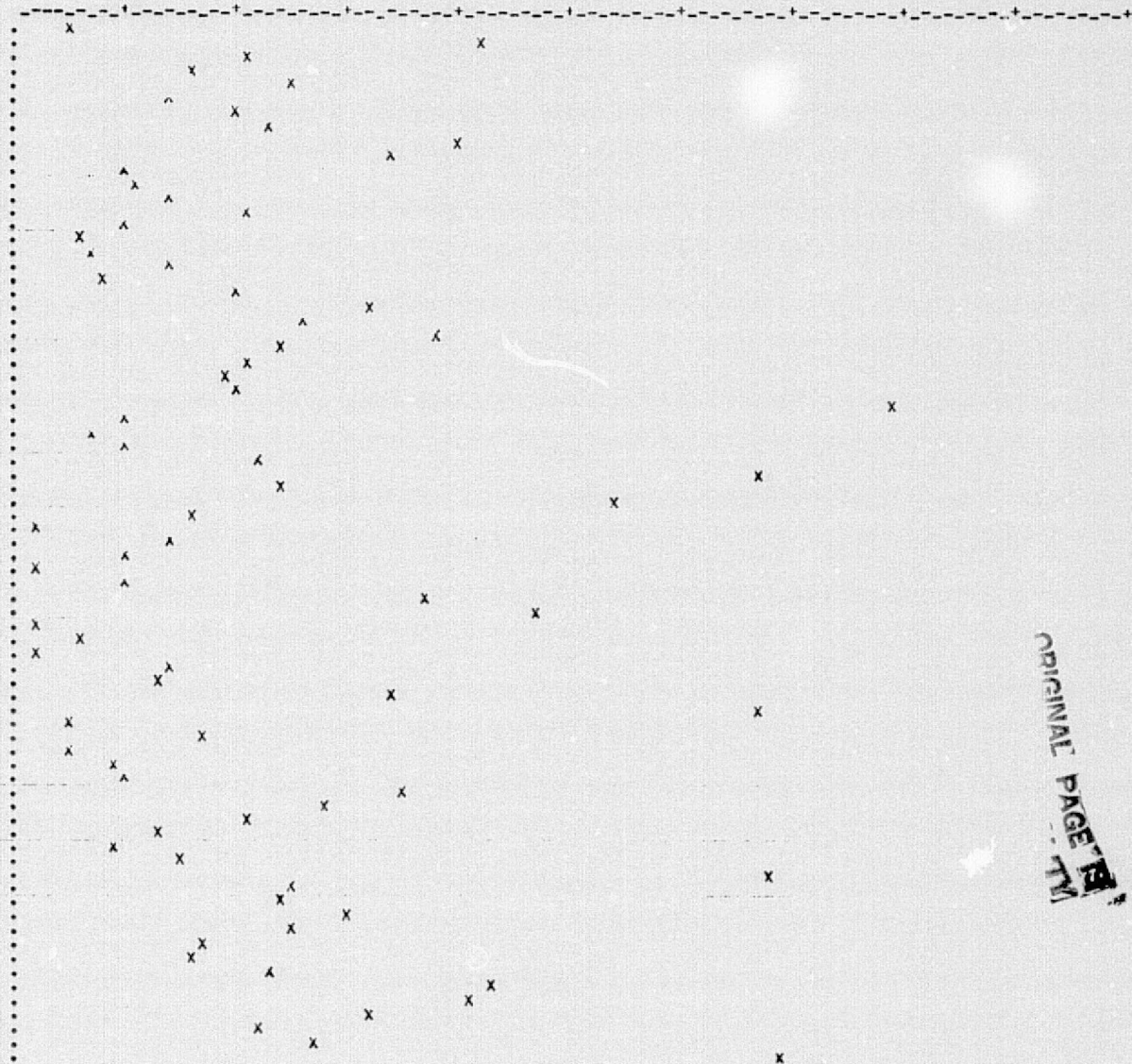
LOWER BOUND OF INTERVAL	UPPER BOUND OF INTERVAL	OBSERVATIONS IN INTERVAL
-9999.	-7.902	2
-7.902	-7.480	2
-7.480	-7.070	3
-7.070	-6.655	3
-6.655	-6.237	1
-6.237	-5.821	3
-5.821	-5.405	8
-5.405	-4.989	11
-4.989	-4.572	14
-4.572	-4.150	30
-4.150	-3.740	33
-3.740	-3.324	23
-3.324	-2.907	9
-2.907	-2.491	3
-2.491	-2.075	0
-2.075	-1.659	0
-1.659	-1.243	0
-1.243	9999.	0

CHI-SQUARED VALUE WITH 40 D. F. IS 05.98

AVERAGE DAILY LEVELS OF SULFUR DIOXIDE FOR QUANTROVE ARMPRY

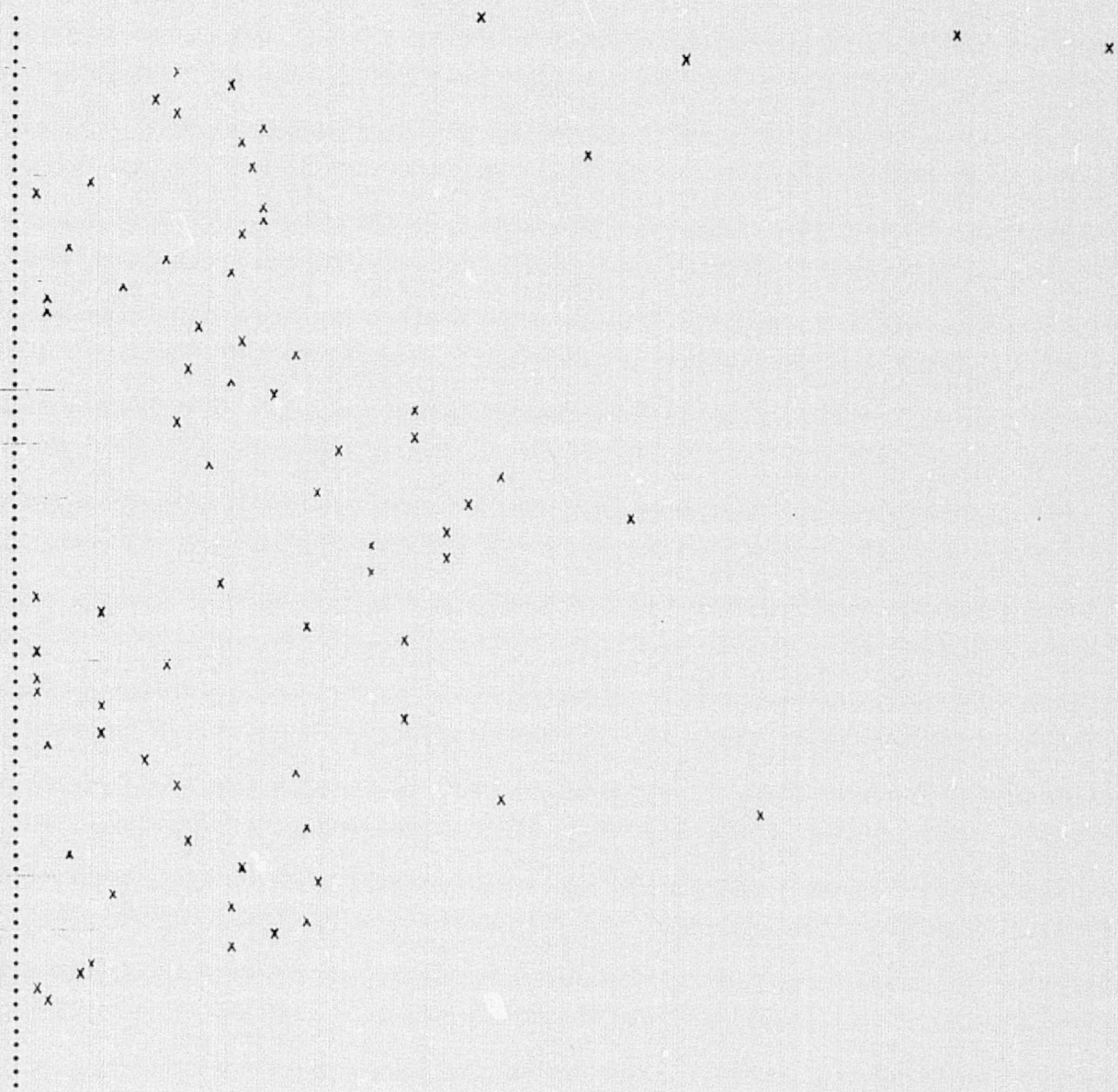
.0076 .0101 .0227 .0303 .0378 .0454 .0530 .0605 .0681 .0757

11 2 72
11 3 72
11 4 72
11 5 72
11 6 72
11 7 72
11 8 72
11 9 72
11 10 72
11 11 72
11 12 72
11 13 72
11 14 72
11 15 72
11 16 72
11 17 72
11 18 72
11 19 72
11 20 72
11 21 72
11 22 72
11 23 72
11 24 72
11 25 72
11 26 72
11 27 72
11 28 72
11 29 72
11 30 72
12 1 72
12 2 72
12 3 72
12 4 72
12 5 72
12 6 72
12 7 72
12 8 72
12 9 72
12 10 72
12 11 72
12 12 72
12 13 72
12 14 72
12 15 72
12 16 72
12 17 72
12 18 72
12 19 72
12 20 72
12 21 72
12 22 72
12 23 72
12 24 72
12 25 72
12 26 72
12 27 72
12 28 72
12 29 72
12 30 72
12 31 72
1 1 73
1 2 73
1 3 73
1 4 73
1 5 73
1 6 73
1 7 73
1 8 73
1 9 73
1 10 73
1 11 73
1 12 73
1 13 73
1 14 73



ORIGINAL PAGE 19

1 15 73
 1 16 73
 1 17 73
 1 18 73
 1 19 73
 1 20 73
 1 21 73
 1 22 73
 1 23 73
 1 24 73
 1 25 73
 1 26 73
 1 27 73
 1 28 73
 1 29 73
 1 30 73
 1 31 73
 2 1 73
 2 2 73
 2 3 73
 2 4 73
 2 5 73
 2 6 73
 2 7 73
 2 8 73
 2 9 73
 2 10 73
 2 11 73
 2 12 73
 2 13 73
 2 14 73
 2 15 73
 2 16 73
 2 17 73
 2 18 73
 2 19 73
 2 20 73
 2 21 73
 2 22 73
 2 23 73
 2 24 73
 2 25 73
 2 26 73
 2 27 73
 2 28 73
 3 1 73
 3 2 73
 3 3 73
 3 4 73
 3 5 73
 3 6 73
 3 7 73
 3 8 73
 3 9 73
 3 10 73
 3 11 73
 3 12 73
 3 13 73
 3 14 73
 3 15 73
 3 16 73
 3 17 73
 3 18 73
 3 19 73
 3 20 73
 3 21 73
 3 22 73
 3 23 73
 3 24 73
 3 25 73
 3 26 73
 3 27 73
 3 28 73
 3 29 73
 3 30 73
 3 31 73
 4 1 73
 4 2 73
 4 3 73
 4 4 73



ORIGINAL PAGE IS
OF POOR QUALITY

A63

5	73
6	73
7	73
8	73
9	73
10	73
11	73
12	73
13	73
14	73
15	73
16	73
17	73
18	73
19	73
20	73
21	73
22	73
23	73
24	73
25	73
26	73
27	73
28	73
29	73
30	73

• • • • •

[illegible]

2.

A66

A67

၇၄၇၅
၇၄၇၆

999
999

ORIGINAL PAGE IS
OF POOR QUALITY

A69

ORIGINAL PAGE IS
OF POOR QUALITY

A71

[illegible]

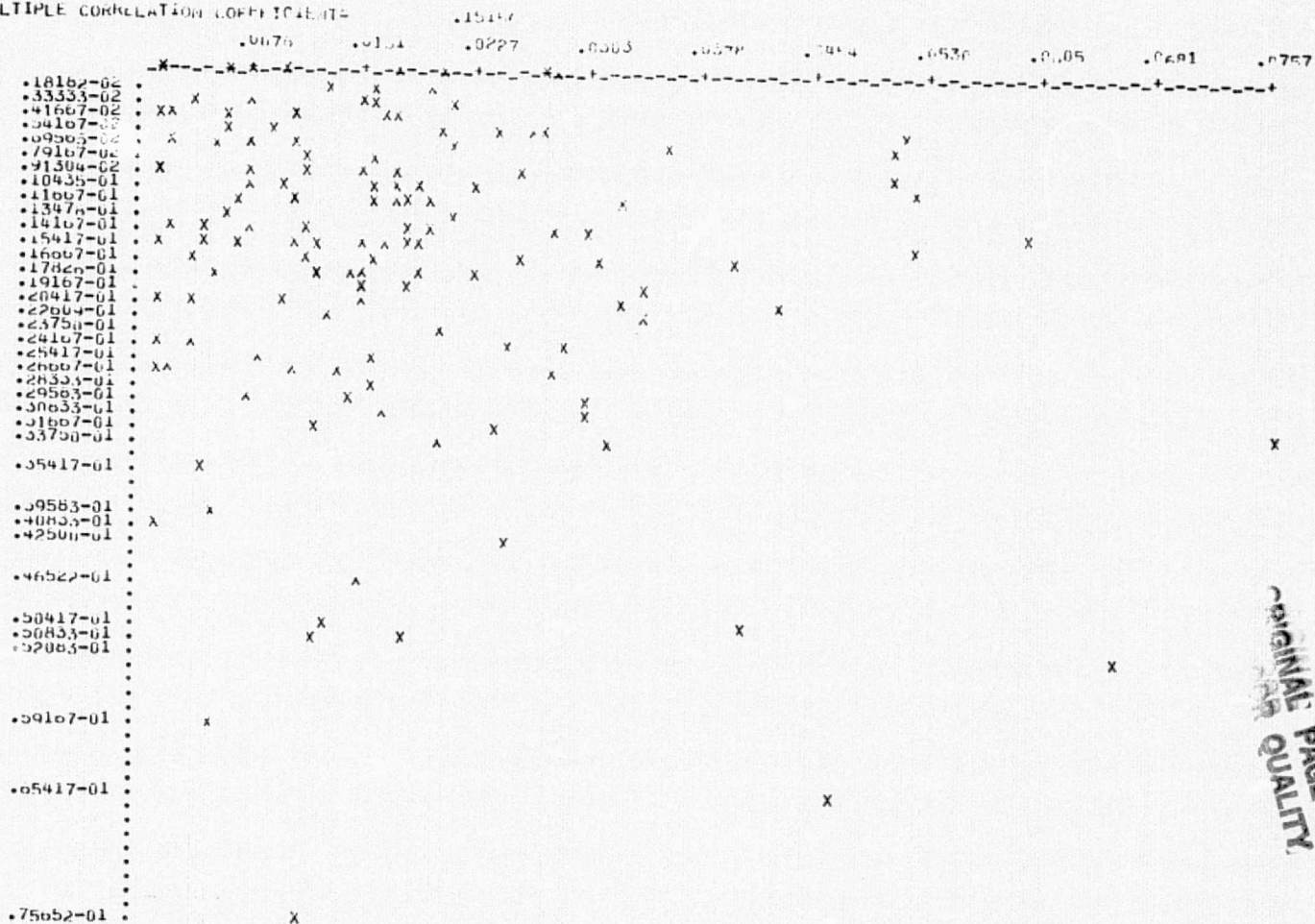
A73

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

X1=INDEPENDENT VARIABLE NO. 11 SULFUR DIOXIDE
 Y=DEPENDENT VARIABLE: SULFUR DIOXIDE CHEMICAL ANALYSIS

REGRESSION EQUATION:
 $Y = .0145 + .1518012 X_1$

MULTIPLE CORRELATION COEFFICIENTS



ORIGINAL PAGE IS
 OF
 QUALITY

Example 5

FOLLOWING ARE DATA RECORDS (Images) in SORT ORDER (Alphabetical)									
440710001F0217211	1	04240114072	1	1	1	1	1	1	0
440710001F0217211	1	1124240114072	1	1	1	1	1	1	0
440710001F0217211	2	04240114072	2	0	0	0	0	0	0
440710001F0217211	2	2124240114072	2	2	2	2	2	2	10
440710001F0217211	3	04240114072	2	1	2	0	1	1	1
440710001F0217211	4	124240114072	2	1	2	2	1	1	1
440710001F0217211	4	04240114072	2	0	0	0	1	1	1
440710001F0217211	5	04240114072	4	1	1	1	1	1	1
440710001F0217211	5	24240114072	4	2	2	2	2	2	1
440710001F0217211	6	04240114072	1	1	1	1	1	1	1
440710001F0217211	6	24240114072	1	2	2	2	2	2	1
440710001F0217211	7	04240114072	2	2	2	2	2	2	1
440710001F0217211	7	124240114072	2	2	2	2	2	2	1
440710001F0217211	8	04240114072	1	2	1	1	1	1	1
440710001F0217211	8	24240114072	2	1	1	1	1	1	1
440710001F0217211	9	04240114072	2	1	1	1	1	1	1
440710001F0217211	9	124240114072	4	4	4	4	4	4	5
440710001F0217211	10	04240114072	3	3	3	3	3	3	4
440710001F0217211	10	24240114072	5	5	5	5	5	5	4
440710001F0217211	11	04240114072	0	0	0	0	0	0	0
440710001F0217211	11	124240114072	1	1	1	1	1	1	1
440710001F0217211	11	24240114072	1	1	1	1	1	1	1
440710001F0217211	12	04240114072	1	1	1	1	1	1	1
440710001F0217211	12	124240114072	1	1	1	1	1	1	1
440710001F0217211	13	04240114072	1	1	1	1	1	1	1
440710001F0217211	13	24240114072	1	2	2	2	2	2	1
440710001F0217211	14	04240114072	1	2	1	1	1	1	1
440710001F0217211	14	124240114072	2	2	2	2	2	2	1
440710001F0217211	15	04240114072	0	0	0	0	0	0	0
440710001F0217211	15	24240114072	1	0	0	0	1	1	1
440710001F0217211	16	04240114072	0	0	0	0	0	0	0
440710001F0217211	16	124240114072	0	0	0	0	0	0	0
440710001F0217211	17	04240114072	1	2	1	0	0	0	0
440710001F0217211	17	24240114072	0	0	0	0	0	0	0
440710001F0217211	18	04240114072	1	1	1	1	1	1	1
440710001F0217211	18	124240114072	1	1	1	1	1	1	1
440710001F0217211	19	04240114072	1	1	1	1	1	1	1
440710001F0217211	19	24240114072	1	1	1	1	1	1	1
440710001F0217211	20	04240114072	1	1	1	1	1	1	1
440710001F0217211	20	124240114072	2	2	2	2	2	2	1
440710001F0217211	21	04240114072	1	1	1	1	1	1	1
440710001F0217211	21	24240114072	1	1	1	1	1	1	1
440710001F0217211	22	04240114072	1	1	1	1	1	1	1
440710001F0217211	22	124240114072	1	1	1	1	1	1	1

ORIGINAL PAGE IS
OF POOR QUALITY

A77

X1: INDEPENDENT VARIABLE NO. 1: TIME WHEN LAST UNIT OF TIME CONSUMED IS 10 MIN
Y: DEPENDENT VARIABLE: SURFACING PERFORMANCE (RANGE 0 TO 100) (COEFFICIENTS 11 TO 17)

REGRESSION EQUATION:
 $Y = .0177 + .0002145 X1$

MULTIPLE CORRELATION COEFFICIENT = .62285

WFIN

ORIGINAL PAGE IS
OF POOR QUALITY